

Chapter 34

Optimization of the Wishart Joint Eigenvalue Probability Density Distribution Based on the Vandermonde Determinant



Asaph Keikara Muhumuza, Karl Lundengård, Jonas Österberg, Sergei Silvestrov, John Magero Mango and Godwin Kakuba

Abstract A number of models from mathematics, physics, probability theory and statistics can be described in terms of Wishart matrices and their eigenvalues. The most prominent example being the Laguerre ensembles of the spectrum of Wishart matrix. We aim to express extreme points of the joint eigenvalue probability density distribution of a Wishart matrix using optimisation techniques for the Vandermonde determinant over certain surfaces implicitly defined by univariate polynomials.

Keywords Vandermonde determinant · Orthogonal ensembles · Gaussian ensembles · Wishart ensembles · Eigenvalue density optimization

MSC 2010 Classification 15A15 · 33C45 · 28C20

A. K. Muhumuza (✉)

Department of Mathematics, Busitema University, Box 236, Tororo, Uganda
e-mail: asaph.keikara.muhumuza@mdh.se

A. K. Muhumuza · K. Lundengård · J. Österberg · S. Silvestrov
Division of Applied Mathematics, School of Education, Culture and Communication,
Mälardalen University, Box 883, 72123 Västerås, Sweden
e-mail: karl.lundengard@mdh.se

S. Silvestrov
e-mail: sergei.silvestrov@mdh.se

J. M. Mango · G. Kakuba
Department of Mathematics, Makerere University, Box 7062, Kampala, Uganda
e-mail: mango@cns.mak.ac.ug

G. Kakuba
e-mail: kakuba@cns.mak.ac.ug

34.1 Introduction

In this work, we review and investigate the Gaussian β -ensembles as the basis for a generalized Wishart density distribution and how this can be optimized over various surfaces, in particular a unit sphere. We take advantage of the general properties of Vandermonde determinant. To begin with, we give a brief outline of key terms including but not limited to Gaussian univariate and multivariate distributions, the Chi-squared density, the Wishart density, the occurrence of random matrices, their joint eigenvalue probability distribution, the β -ensembles, the Vandermonde matrix and its determinant. We then illustrate the optimization of the joint probability density function of the β -ensembles over a unit sphere based on the characteristic properties of the Vandermonde determinant.

34.1.1 Univariate and Multivariate Normal Distribution

Definition 34.1 The univariate normal probability density function (Gaussian normal density) for a random variable X , which is the basis for construction of many multivariate distributions that occur in statistics, can be expressed as [4]:

$$P_X(x) = k \exp \left\{ -\frac{1}{2} \alpha (x - \beta)^2 \right\} \equiv k \exp \left\{ -\frac{1}{2} (x - \beta) \alpha (x - \beta) \right\} \tag{34.1}$$

where α and k is chosen so that the integral of (34.1) over the entire x -axis is unity and β is equal to the expectation of X , that is, $E[X] = \beta$. It is then said that X follows a normal probability density function with parameters α and β , also expressed as $X \sim \mathcal{N}(\alpha, \beta)$.

The density function of the multivariate normal distribution of random variables say X_1, \dots, X_p is defined analogously. If the scalar variable x in (34.1) is directly replaced by the vector $\mathbf{X} = (X_1, \dots, X_p)^T$, the scalar constant β is replaced by a vector $\mathbf{b} = (b_1, \dots, b_p)^T$ and the positive definite matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{pmatrix}. \tag{34.2}$$

The expression

$$\alpha(x - \beta)^2 = (x - \beta) \alpha (x - \beta)$$

is replaced by the quadratic form

$$(\mathbf{X} - \mathbf{b})^\top \mathbf{A}(\mathbf{X} - \mathbf{b}) = \sum_{i,j=1}^p a_{ij}(x_i - b_i)(x_j - b_j). \tag{34.3}$$

Thus, the density of the p -variate normal distribution becomes

$$\mathbb{P}(\mathbf{X}) = K \exp \left\{ \frac{1}{2}(\mathbf{X} - \mathbf{b})^\top \mathbf{A}(\mathbf{X} - \mathbf{b}) \right\} \tag{34.4}$$

where \top denotes transpose and $K > 0$ is chosen so that the integral over the entire p -dimensional Euclidean space x_1, \dots, x_p is unity.

Theorem 34.1 *If the density of a p -dimensional random vector \mathbf{X} is*

$$\sqrt{|\mathbf{A}|}(2\pi)^{-\frac{1}{2}p} \exp \left\{ -\frac{1}{2}(\mathbf{X} - \mathbf{b})^\top \mathbf{A}(\mathbf{X} - \mathbf{b}) \right\},$$

then the expected value of \mathbf{X} is \mathbf{b} and the covariance matrix is \mathbf{A}^{-1} , see [4]. Conversely, given a vector $\boldsymbol{\mu}$ and a positive definite matrix $\boldsymbol{\Sigma}$, there is a multivariate normal density

$$\mathbb{P}(\mathbf{X}) = (2\pi)^{-\frac{1}{2}p} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left\{ (\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \right\} \tag{34.5}$$

such that the expected value of the density is $\boldsymbol{\mu}$ and the covariance matrix is $\boldsymbol{\Sigma}$.

The density (34.5) is often denoted as $\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

For example, the diagonal elements of the covariance matrix, $\boldsymbol{\Sigma}_{ii}$, is the variance of the i th component of \mathbf{X} , which may sometimes be denoted by σ_i^2 . The correlation between X_i and X_j is defined as

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} = \frac{\sigma_{ij}}{\sigma_i\sigma_j}$$

where σ_k denotes the standard deviation of X_k and $\sigma_{ij} = \boldsymbol{\Sigma}_{ij}$. This measure of association is symmetric in X_i and X_j such that $\rho_{ij} = \rho_{ji}$. Since

$$\begin{pmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{pmatrix} = \begin{pmatrix} \sigma_i^2 & \sigma_i\sigma_j\rho_{ij} \\ \sigma_i\sigma_j\rho_{ij} & \sigma_j^2 \end{pmatrix}$$

is positive-definite, the determinant

$$\begin{vmatrix} \sigma_i^2 & \sigma_i\sigma_j\rho_{ij} \\ \sigma_i\sigma_j\rho_{ij} & \sigma_j^2 \end{vmatrix} = \sigma_i^1\sigma_j^2(1 - \rho_{ij}^2)$$

is positive. Therefore $-1 < \rho_{ij} < 1$.

34.1.2 Wishart Distribution

The matrix distribution that is now known as a Wishart distribution, was first derived by Wishart in the late 1920s [56]. It is usually regarded as a multivariate extension of the χ^2 -distribution.

Theorem 34.2 *The sum of squares, $\chi^2 = Z_1^2 + \dots + Z_n^2$ of n -independent standard normal variables Z_i of mean 0 and variance 1, that is, distributed as $\mathcal{N}(0, 1)$ has a χ^2 -distribution defined by:*

$$\mathbb{P}_{\chi^2}(x) = \frac{1}{2^{\frac{1}{2}n} \Gamma(\frac{1}{2}n)} e^{-\frac{1}{2}x^2} (\chi^2)^{\frac{1}{2}n-1}. \tag{34.6}$$

where $\Gamma(\cdot)$ is the Gamma function [40].

Definition 34.2 Let $\mathbf{X} = (X_1, \dots, X_n)$, where $X_i \sim \mathcal{N}(\mu_i, \Sigma)$ and \mathbf{X}_i is independent of \mathbf{X}_j , where $i \neq j$. The matrix $\mathbf{W} : p \times p$ is said to be Wishart distributed [56] if and only if $\mathbf{W} = \mathbf{X}\mathbf{X}^T$ for some matrix \mathbf{X} in a family of Gaussian matrices $\mathbf{G}_{m \times n}$, $m \leq n$, that is, $\mathbf{X} \sim \mathcal{N}_{m,n}(\boldsymbol{\mu}, \Sigma, \mathbf{I})$ where $\Sigma \geq 0$. If $\boldsymbol{\mu} = 0$ we have a central Wishart distribution which will be denoted by $\mathbf{W} \sim \mathcal{W}_m(\Sigma, n)$, and if $\boldsymbol{\mu} \neq 0$ we have a non-central Wishart distribution which will be denoted $\mathbf{W} \sim \mathcal{W}_m(\Sigma, n, \Delta)$, where $\Delta = \boldsymbol{\mu}\boldsymbol{\mu}^T$ and n is the number of degrees of freedom.

In our study, we shall mainly focus on the central Wishart distribution for which $\boldsymbol{\mu} = 0$ and $\mathbf{X} \sim \mathcal{N}_{m,n}(\boldsymbol{\mu}, \Sigma, \mathbf{I})$

Theorem 34.3 ([4]) *Given a random matrix \mathbf{W} which can be expressed as $\mathbf{W} = \mathbf{X}\mathbf{X}^T$ where $\mathbf{X}_1, \dots, \mathbf{X}_n$, ($n \geq p$) are independent, each with the distribution $\mathcal{N}_p(\boldsymbol{\mu}, \Sigma)$. Then, the distribution of $\mathbf{W} \sim \mathcal{W}_p(\Sigma, n)$. If $\Sigma > 0$, then the random matrix \mathbf{W} has a joint density functions:*

$$\mathbb{P}(\mathbf{W}) = \begin{cases} \frac{1}{2^{\frac{np}{2}} \Gamma_p(\frac{n}{2})} |\mathbf{W}|^{\frac{n-p-1}{2}} \exp\left(-\frac{1}{2}\text{Tr}(\Sigma^{-1}\mathbf{W})\right), & \text{if } \mathbf{W} > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{34.7}$$

where the multivariate Gamma function is given by

$$\Gamma_p(n/2) = \pi^{\frac{p(n-1)}{2}} \prod_{i=1}^p \Gamma\left(\frac{1}{2}(n+1-i)\right). \tag{34.8}$$

If $p = 1$, $\boldsymbol{\mu} = \mathbf{0}$ and $\Sigma = \mathbf{1}$, then the Wishart matrix is identical to a central χ^2 -variable with n degrees of freedom as defined in (34.6).

Theorem 34.4 ([24, 39]) *If \mathbf{X} is distributed as $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then the probability density distribution of the eigenvalues of $\mathbf{X}\mathbf{X}^\top$, denoted $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)$, is given by:*

$$\mathbb{P}(\boldsymbol{\lambda}) = \frac{\pi^{-\frac{1}{2}n} \det(\boldsymbol{\Sigma})^{-\frac{1}{2}n} \det(\mathbf{D})^{\frac{1}{2}(n-p-1)}}{2^{\frac{1}{2}np} \Gamma_p(\frac{1}{2}n) \Gamma_p(\frac{1}{2}p)} \prod_{i < j} (\lambda_i - \lambda_j) \exp\left(-\frac{1}{2} \text{Tr}(\boldsymbol{\Sigma}^{-1} \mathbf{D})\right) \tag{34.9}$$

where $\mathbf{D} = \text{diag}(\lambda_i)$ and Γ is the Gamma function.

It will prove useful that (34.9) contains the term $\prod_{i < j} (\lambda_i - \lambda_j)$ which is the determinant of a Vandermonde matrix [46]. A Vandermonde matrix is a well-known type of matrix that appears in many different applications both in mathematics, physics and recently in multivariate statistics, most famously curve-fitting using polynomials, for details see [46].

Definition 34.3 Square Vandermonde matrices of size $n \times n$ are determined by N values $\mathbf{x} = (x_1, \dots, x_n)$ and is defined as follows:

$$V_n(\mathbf{x}) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{bmatrix}. \tag{34.10}$$

The determinant of the Vandermonde matrix is well known.

Lemma 34.1 *The determinant of square Vandermonde matrices has the form*

$$\det V_n(\mathbf{x}) \equiv v_n(\mathbf{x}) = \prod_{1 \leq i < j \leq n} (x_j - x_i). \tag{34.11}$$

This determinant is also referred to as the *Vandermonde determinant* or *Vandermonde polynomial* or *Vandermondian* [46].

We take advantage of this fact of Vandermonde determinant to establish the relationship between the product of Vandermonde matrices and joint eigenvalue probability density functions for large random matrices that occur in various areas of both classical mechanics, mathematics, statistics and many other areas of science. We also illustrate the optimization of these densities based of the extreme points Vandermonde determinant.

The extreme points of the Vandermonde determinant appears in random matrix theory, for example to compute the limiting value of the so called Stieltjes transform using the method sometimes called the ‘Coulomb gas analogy’ [32]. This is also closely related to many problems in quantum mechanics and statistical mechanics. For an overview of some other applications of the extreme points see [35].

In the next section we give a brief overview of random matrix theory (RMT).

34.2 Overview of Random Matrix Theory

Random matrices were first introduced in mathematical statistics in the late 1920s [56] and today the joint probability density function of eigenvalues of random matrices play a significant role both in probability theory, mathematical physics and quantum mechanics [22]. A random matrix, in simple terms can be defined as any matrix whose real or complex valued entries are random variables.

Random matrix theory primarily discusses the properties large or complex matrices with random variables as entries by utilizing the existing probability laws, in particular, Gaussian distributions [4, 5]. The main motivational question in the probabilistic approach to random matrices is: what can be said about the probabilities of a few or if not all of its eigenvalues and eigenvectors? This question is significant in many areas of science including particle physics, mathematics, statistics and finance as highlighted here under.

In nuclear physics random matrices were applied in the modelling of the nuclei of heavy atoms [55]. The main idea was to investigate the spacing between the lines in the electromagnetic spectrum of a heavy atom nucleus, e.g. Uranium 238, which resembles the separation between the eigenvalues of a random matrix [32]. These random matrices have also been employed in solid-state physics to model the chaotic behaviour of large disordered Hamiltonians in terms of mean field approximation [14]. Random matrices have also been applied in quantum chaos to characterise the spectral statistics of quantum systems [9, 12].

Random unitary matrix transformations has also appears in theoretical physics, e.g. the boson sampling model [1] has been applied in quantum optics to describe the advantages of quantum computation over classical computation. Random unitary transformations can also be directly implemented in an optical circuit, by mapping their parameters to optical circuit components [41].

Other applications in theoretical physics include, analysing the chiral Dirac operator [28, 52] quantum chromodynamics, quantum gravity in two dimensions [21], in mesoscopic physics random matrices are used to characterise materials of intermediate length [43], spin-transfer torque [42], the fractional quantum Hall effect [10], Anderson localization [25], quantum dots [59] and superconductors [7], electrodynamic properties of structural materials [58], describing electrical conduction properties of disordered organic and inorganic materials [57], quantum gravity [15] and string theory [8].

In mathematics some application include the distribution of the zeros of the Riemann zeta function [27], enumeration of permutations having certain particularities in which the random matrices can help to derive polynomials permutation patterns [38], counting of certain knots and links as applies to folding and coloring [8].

In multivariate statistics random matrices were introduced for statistical analysis of large samples in estimation of covariance matrices [18–20, 33, 45, 56]. More significant results have proven that to extend the classical scalar inequalities for improved analysis of a structured dimension reduction based on largest eigenvalues of finite sums of random Hermitian matrices [48].

Random matrices have also been applied to financial modelling especially risk models and time series [6, 23, 51, 56].

Random matrices also are increasingly used to model the network of synaptic connections between neurons in the brain as applies to neural networks or neuroscience. Neuronal networks can help to construct dynamical models based on random connectivity matrix [44]. This has also helped to establish the link relating the statistical properties of the spectrum of biologically inspired random matrix models to the dynamical behaviour of randomly connected neural networks [13, 26, 36, 47, 53].

In optimal control theory random matrices appear as coefficients in the state equation of linear evolution. In most problems the values of the parameters in these matrices are not known with certainty, in which case there are random matrices in the state equation and the problem is known as one of stochastic control [11, 49, 50].

In the next section, we will discuss some well-known ensembles that that appear in the mathematical study of random matrices.

34.3 Classical Random Matrix Ensembles

The key famously known classical ensembles include the Gaussian Orthogonal Ensembles (G.O.E), the Gaussian Unitary Ensembles (G.U.E), the Gaussian Symplectic Ensembles (GSE), the Wishart Ensembles (W.E), the MANOVA Ensembles (M.E) and the Circular Ensembles (C.E). These can be derived from the multivariate Gaussian matrix, \mathbf{G}_β , $\beta = 1, 2, 4$. Since, the multivariate Gaussian possesses an inherent orthogonal property from the standard normal distribution, that is, they remain invariant under orthogonal transformations. More detailed discussions on these ensembles can be found in [3, 4, 32, 37, 54, 56].

Definition 34.4 ([29]) The Gaussian Orthogonal Ensembles (G.O.E) are characterised by the symmetric matrix $\mathbf{X} = \mathbf{G}_1(N, N)$ obtained as $(\mathbf{X} + \mathbf{X}^\top)/2$. The diagonal entries of \mathbf{X} are independent and identically distributes (i.i.d) with a standard normal distribution $\mathcal{N}(0, 1)$ while the off-diagonal entries are i.i.d with a standard normal distribution $\mathcal{N}_1(0, 1/2)$. That is, a random matrix \mathbf{X} is called the Gaussian Orthogonal Ensemble (GOE), if it is symmetric and real-valued ($X_{ij} = X_{ji}$) and has

$$X_{-ij} = \begin{cases} \sqrt{2}\xi_{ii} \sim \mathcal{N}_1(0, 1), & \text{if } i = j \\ \xi_{ij} \sim \mathcal{N}_1(0, 1/2), & i < j. \end{cases} \tag{34.12}$$

Definition 34.5 ([16, 29]) The Gaussian Unitary Ensembles (G.U.E), are characterised by the Hermitian complex-valued matrix $\mathbf{H} = \mathbf{G}_2(N, N)$ obtained as $(\mathbf{H} + \mathbf{H}^{\top*})/2$ where \top^* is the operation of taking the Hermitian transpose, that is, the Hermitian or conjugate transpose of \mathbf{H} , and expressed as $(\mathbf{H}^{\top*})_{ij} = \overline{\mathbf{H}}_{ji}$. The diagonal entries of \mathbf{H} are independent and identically distributes (i.i.d) with a standard normal distribution $\mathcal{N}(0, 1)$ while the off-diagonal entries are i.i.d with a

standard normal distribution $\mathcal{N}_2(0, 1/2)$. That is, random matrix \mathbf{H} is called a Gaussian Unitary Ensemble (GUE), if it is complex-valued, Hermitian ($\mathbf{H}_{ij}^{\top*} = \overline{\mathbf{H}}_{ji}$), and the entries satisfy

$$\mathbf{H}_{ij} = \begin{cases} \sqrt{2}\xi_{ii} \sim \mathcal{N}_2(0, 1), & \text{if } i = j \\ \frac{1}{\sqrt{2}}(\xi_{ij} + \sqrt{-1}\eta_{ij}) \sim \mathcal{N}_2(0, 1/2), & i < j. \end{cases} \tag{34.13}$$

Definition 34.6 ([6, 29]) The Gaussian Symplectic Ensembles (GSE), are characterised by the self-dual matrix $\mathbf{S} = \mathbf{G}_4(N, N)$ obtained as $(\mathbf{S} + \mathbf{S}^{\top*})/2$ where \top^* represents the operation of taking the conjugate transpose of a quaternion matrix. The diagonal entries \mathbf{H} are independent and identically distributed (i.i.d) with a standard normal distribution $\mathcal{N}(0, 1)$ while the off-diagonal entries are i.i.d with a standard normal distribution $\mathcal{N}_4(0, 1/2)$.

Definition 34.7 ([6, 29]) The Wishart Ensembles (W.E), $\mathcal{W}_\beta(m, n)$, $m \geq n$, are characterised by the symmetric, Hermitian or self-dual matrix $\mathbf{W} = \mathbf{W}_\beta(N, N)$ obtained as $\mathbf{W} = \mathbf{A}\mathbf{A}^\top$, $\mathbf{W} = \mathbf{H}\mathbf{H}^\top$, or $\mathbf{W} = \mathbf{S}\mathbf{S}^\top$ where \top represents the operation of taking the usual transposes of defined in G.O.E, G.U.E and G.S.E above respectively.

Definition 34.8 ([6, 29]) The MANOVA Ensembles (M.E), $\mathcal{J}_\beta(m_1, m_2, n)$, $m_1, m_2 \geq n$, are characterised by the symmetric, Hermitian or self-dual matrix $\mathbf{A}/(\mathbf{A} + \mathbf{B})$ where \mathbf{A} and \mathbf{B} are $\mathbf{W}_\beta(m_1, n)$ and $\mathbf{W}_\beta(m_2, n)$ respectively.

Definition 34.9 ([16, 29]) The Circular Ensembles (C.E), are characterised by the special matrix $\mathbf{U}\mathbf{U}^\top$ where \mathbf{U}_β , $\beta = 1, 2$ is a uniformly distributed unitary matrix.

Lemma 34.2 ([29]) From the Gaussian normal distribution with mean μ and variance σ^2 , that is, $\mathbf{X} \sim \mathcal{N}(\mu, \sigma^2)$, given by (34.1) and the multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and the covariance matrix is $\boldsymbol{\Sigma}$, $\mathcal{N}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ given in (34.5), then it can be verified that the joint density of \mathbf{A} is written as:

$$\mathbb{P}_{\mathbf{X}}(\mathbf{A}) = \frac{1}{2^{n/2}} \frac{1}{\pi^{n(n+1)/4}} \exp(-\|\mathbf{A}\|_F^2/2)$$

where $\|\mathbf{A}\|_F$ represents the Frobenius norm of \mathbf{A} .

Theorem 34.5 ([4, 16]) If we let \mathbf{X} be an $N \times N$ random matrix with entries that are independently identically distributed as $\mathcal{N}(0, 1)$, then the joint density distribution of the Gaussian ensembles is given by:

Gaussian: $\begin{cases} \text{Orthogonal } \beta = 1 \\ \text{Unitary } \beta = 2 \\ \text{Symplectic } \beta = 4 \end{cases} \mathbb{P}_\beta(\mathbf{A}) = \frac{1}{2^{n/2}} \frac{1}{\pi^{n(n+1) \beta/4}} \exp\left(-\frac{1}{2}\|\mathbf{A}\|_F^2\right).$

Theorem 34.6 ([16, 32]) Considering a Wishart matrix $\mathbf{W}_\beta(m, n) = \mathbf{X}\mathbf{X}^\top$ where $\mathbf{X} = \mathbf{G}_\beta(m, n)$ is a multivariate Gaussian matrix. Then, the joint elements of

$W_\beta(m, n)$ can be computed in two steps, first writing $W = QR$ and then integrating out Q leaving R . Secondly applying the transformation $W = RR^T$, which is the famous Cholesky factorization of matrices in numerical analysis. Then the joint density distribution for Wishart ensembles of W is given by:

$$\text{Wishart: } \begin{cases} \text{Orthogonal } \beta = 1 \\ \text{Unitary } \beta = 2 \\ \text{Symplectic } \beta = 4 \end{cases} \quad \mathbb{P}_\beta(W) = \frac{\exp(-\text{tr}(W/2)) (\det W)^{\beta(m-n+1)/2-1}}{2^{mn\beta/2} \Gamma_n^\beta(m\beta/2)}.$$

Here we notice that the density distribution for both the Gaussian and Wishart ensembles are made up of determinant term and exponential trace term. This generalizes the fact that indeed the determinant term is actually the Vandermonde determinant in (34.11) for the case of the joint eigenvalue density functions. This concept further explained in the next section.

34.4 The Vandermonde Determinant and Joint Eigenvalue Probability Densities for Random Matrices

To obtain the joint eigenvalue densities for random matrices, we apply the the principle of matrix factorization, for instance if the random matrix X is expressed as $X = Q\Lambda Q^T$, then Λ directly gives the eigenvalues X [24]. Applying the Jacobian technique for joint density transformation, see for example [4], this yields the joint densities of eigenvalues and eigenvectors.

Lemma 34.3 *The three Gaussian ensembles have joint eigenvalues probability density function [32, 37] given by*

$$\text{Gaussian: } \mathbb{P}_\beta(\lambda) = C_N^\beta \prod_{i < j} |\lambda_1 - \lambda_2|^\beta \exp\left(-\frac{1}{2} \sum_{i=1}^N \lambda_i^2\right) \tag{34.14}$$

where $\beta = 1$ representing reals, $\beta = 2$ representing the complexes, and $\beta = 4$ representing the quaternion, and

$$C_N^\beta = (2\pi)^{-N/2} \prod_{j=1}^N \frac{\Gamma(1 + \beta/2)}{\Gamma(1 + j\beta/2)}.$$

Lemma 34.4 ([24, 32]) *The Wishart (or Laguerre) ensembles have a joint eigenvalue probability density distribution given by*

$$\text{Wishart: } \mathbb{P}_\beta(\lambda) = C_N^{\beta,\alpha} \prod_{i < j} |\lambda_1 - \lambda_2|^\beta \prod_i \lambda_i^{\alpha-p} \exp\left(-\frac{1}{2} \sum_{i=1}^N \lambda_i^2\right) \tag{34.15}$$

where $\alpha = \frac{\beta}{2}m$ and $p = 1 + \frac{\beta}{2}(N - 1)$. The β parameter is decided by what type of elements are in the Wishart matrix, real-valued elements corresponds to $\beta = 1$, complex-valued elements correspond to $\beta = 2$ and quaternions elements correspond to $\beta = 4$, and the normalizing constant $C_N^{\beta,\alpha}$ is given by

$$C_N^{\beta,\alpha} = 2^{-N\alpha} \prod_{j=1}^N \frac{\Gamma(1 + \beta/2)}{\Gamma(1 + j\beta/2) \Gamma(\alpha - \frac{\beta}{2}(n - j))} \tag{34.16}$$

Thus the joint eigenvalue probability density distribution for all the ensembles can be summarized in the following theorem [18, 29, 32].

Theorem 34.7 Suppose that $\mathbf{X}_N \in \mathcal{H}^\beta$ for $\beta = 1, 2, 4$. Then, the distribution of eigenvalues of \mathbf{X}_N is given by

$$\mathbb{P}_{\mathbf{X}}(x_1, \dots, x_N) = \bar{C}_N^\beta \prod_{i < j} |x_i - x_j|^\beta \exp\left(-\frac{\beta}{4} \sum_i x_i^2\right) \tag{34.17}$$

where $\bar{C}_N^{(\beta)}$ are normalized constants and can be computed explicitly.

From (34.17) it should be noted that trivially, the properties of a probability density function, that is,

$$0 \leq \mathbb{P}(\mathbf{x}) \leq 1 \text{ and } \int_{\mathbb{R}^N} \mathbb{P}(\mathbf{x}) \prod_{i=1}^N dx_i = 1$$

do hold as verified in [32]. We also notice that the term $\prod_{i < j} |x_i - x_j|^\beta$ in the expression (34.17) is the determinant of the famous Vandermonde matrix (34.10) raised to the power $\beta = 1, 2, 4$. For example, from (34.10) and (34.11) and applying the principles of linear algebra, that is, if \mathbf{A} is an $N \times N$ matrix, then $|\mathbf{A}^\beta| = |\mathbf{A}|^\beta$, for determinants. Thus,

$$|V_N(\mathbf{x})|^\beta = \left| \begin{matrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_N \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{N-1} & x_2^{N-1} & \dots & x_N^{N-1} \end{matrix} \right|^\beta = \left| \prod_{1 \leq i < j \leq N} (x_j - x_i) \right|^\beta = \prod_{1 \leq i < j \leq N} |x_j - x_i|^\beta \tag{34.18}$$

It should also be noted that the trace exponential term

$$\exp\left(-\frac{\beta}{4} \sum_i x_i^2\right)$$

in (34.17) is a product of weight functions of Freud type [30, 34] of the form

$$\omega(x) = \exp(-\alpha x^2),$$

where $\alpha = 1, 1/2, 1/4$. The weights $\omega(x)$ are bounded, that is, $0 \leq |\omega(x)| \leq 1$. If we assume the random variables $\mathbf{X} = \{x_1, \dots, x_N\}$ having normal probability density function with mean $\mu = 0$ and variance $\sigma^2 = 1$, that is, x_i are independent identically distributed i.i.d. as $\mathcal{N}(0, 1)$, then it follows that we can construct the normal density in terms of a Gaussian weights such that

$$\mathbb{P}_{\mathbf{X}}(x_i) = \frac{1}{\sqrt{2\pi}}\omega(x_i).$$

Also, from the definition of the p -th moment of a probability density function

$$\mathbb{E}[X^p] = \int x_i^p \mathbb{P}_{\mathbf{X}}(x_i) dx_i$$

and we have equivalently in terms of Gaussian weights

$$\mathbb{E}[X^p] = \int x_i^p \mathbb{P}_{\mathbf{X}}(x_i) dx_i = \int x_i^p \cdot \frac{1}{\sqrt{2\pi}}\omega(x_i) dx_i.$$

Thus focusing on the coefficients terms of x_i , that is,

$$x_i^p \cdot \frac{1}{\sqrt{2\pi}}\omega(x_i) = kx_i^p \omega(x_i), \quad k = 1/\sqrt{2\pi},$$

we generate a weighted Vandermonde matrix of the $\omega(\mathbf{x})$ weighted form as follows:

$$V_N(\omega(\mathbf{x}\mathbf{x})) = \begin{bmatrix} \omega_1(x_1) & \omega_2(x_1) & \cdots & \omega_N(x_1) \\ \omega_1(x_1)x_1 & \omega_2(x_2)x_2 & \cdots & \omega_N(x_N)x_N \\ \omega_1(x_1)x_1^2 & \omega_2(x_2)x_2^2 & \cdots & \omega_N(x_N)x_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ \omega_1(x_1)x_1^{N-1} & \omega_2(x_2)x_2^{N-1} & \cdots & \omega_N(x_N)x_N^{N-1} \end{bmatrix}. \quad (34.19)$$

The determinant of the Vandermonde matrix in (34.19) can also be obtained taking advantage of the Gaussian weights of the form:

$$\omega_i = \omega(x_i) = \exp\left(-\frac{1}{4}x_i^2\right)$$

and properties of determinant that is, if say \mathbf{A} is an $N \times N$ matrix, then $|\alpha\mathbf{A}| = \alpha^N|A|$. Thus,

$$\begin{aligned}
 \left| \omega(\mathbf{x}) V_N(\mathbf{x}) \right|^\beta &= \left| \begin{matrix} \omega_1(x_1) & \omega_2(x_1) & \cdots & \omega_N(x_1) \\ \omega_1(x_1)x_1 & \omega_2(x_2)x_2 & \cdots & \omega_N(x_N)x_N \\ \omega_1(x_1)x_1^2 & \omega_2(x_2)x_2^2 & \cdots & \omega_N(x_N)x_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ \omega_1(x_1)x_1^{N-1} & \omega_2(x_2)x_2^{N-1} & \cdots & \omega_N(x_N)x_N^{N-1} \end{matrix} \right|^\beta \\
 &= \left| \begin{matrix} \omega_1(x_1) & 0 & 0 & \cdots & 0 \\ 0 & \omega_2(x_2) & 0 & \cdots & 0 \\ 0 & 0 & \omega_3(x_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \omega_N(x_N) \end{matrix} \right| \left| \begin{matrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_N \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_N^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{N-1} & x_2^{N-1} & x_3^{N-1} & \cdots & x_N^{N-1} \end{matrix} \right|^\beta \tag{34.20} \\
 &= \left| \prod_{i=1}^N \omega(x_i) \prod_{1 \leq i < j \leq N} (x_j - x_i) \right|^\beta \\
 &= \left[\exp\left(-\frac{1}{4}x_1^2\right) \cdots \exp\left(-\frac{1}{4}x_N^2\right) \right]^\beta \left| \prod_{1 \leq i < j \leq N} (x_j - x_i) \right|^\beta \\
 &= \left[\exp\left(-\frac{1}{4} \sum_{i=1}^N x_i^2\right) \right]^\beta \prod_{1 \leq i < j \leq N} |x_j - x_i|^\beta \\
 &= \prod_{1 \leq i < j \leq N} |x_j - x_i|^\beta \exp\left(-\frac{\beta}{4} \sum_{i=1}^N x_i^2\right).
 \end{aligned}$$

If \mathbf{X} has a central normal distribution, then for any finite non-negative integer p the plain central moments are given by

$$\mathbb{E}[\mathbf{X}^p] = \begin{cases} 0, & \text{if } p \text{ is odd} \\ \sigma^p (p - 1)!!, & \text{if } p \text{ is even.} \end{cases} \tag{34.21}$$

where $n!!$ denotes the double factorial, that is, the product of numbers from n to 1 that have same parity as n .

The absolute central moment coincides with the plain moments for all even orders and are non-zero for odd orders. Thus, for any non-negative integer p

$$\mathbb{E}[|\mathbf{X}|^p] = \sigma^p (p - 1)!! \cdot \begin{cases} \sqrt{\frac{2}{\pi}}, & \text{if } p \text{ is odd} \\ 1, & \text{if } p \text{ is even} \end{cases} = \sigma^p \cdot \frac{2^{p/2} \Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}. \tag{34.22}$$

Considering $\mathbf{X}_1, \dots, \mathbf{X}_N$ independently normally distributed random variables with mean $\mu = 0$, then the p -th product moment can be expressed as

$$\mathbb{E}[(\mathbf{X}_1 \cdots \mathbf{X}_N)^p] = \mathbb{E}[\mathbf{X}_1^p] \cdots \mathbb{E}[\mathbf{X}_N^p].$$

Thus from (34.22), the joint p -th product moment will be given by

$$\mathbb{E} [|\mathbf{X}_1|^p] \cdots \mathbb{E} [|\mathbf{X}_N|^p] = \left[\sigma^p (p-1)!! \cdot \begin{cases} \sqrt{\frac{2}{\pi}}, & \text{if } p \text{ is odd} \\ 1, & \text{if } p \text{ is even} \end{cases} \right]^N = \left[\sigma^p \cdot \frac{2^{p/2} \Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}} \right]^N.$$

This, thorough examination, generates an equivalent expression for multivariate Gamma function as defined in (34.16), which is also the normalizing constant for the joint eigenvalue density for β -ensembles as given in (34.17). Thus, the same normalizing coefficient can be introduced in the expression (34.20) to lead to the same result as in (34.17).

Basing on the above close link between the Vandermonde determinant, then it is plausible enough to consider the general optimization of Vandermonde determinant over the polynomial constraint defined by trace factor $\sum_{i=1}^N x_i^2$ in the bounded exponential term. We will apply the method of Lagrange multipliers to optimize the density (34.12) to optimize the Vandermonde determinant on the unit sphere and other surfaces, which in turn optimize the joint eigenvalue density as will be demonstrated in the next section.

34.5 Optimising the Joint Eigenvalue Probability Density Function

Lemma 34.5 *For any symmetric $n \times n$ matrix A with eigenvalues $\{\lambda_i, i = 1, \dots, n\}$ that are all distinct, and any polynomial P :*

$$\sum_{k=1}^n P(\lambda_k) = \text{Tr} (P(\mathbf{A})) .$$

Proof By definition, for any eigenvalue λ and eigenvector \mathbf{v} we must have $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ and thus

$$P(\mathbf{A})\mathbf{v} = \left(\sum_{k=0}^m c_k \mathbf{A}^k \right) \mathbf{v} = \sum_{k=0}^m c_k (\mathbf{A}^k \mathbf{v}) = \sum_{k=0}^m c_k \lambda^k \mathbf{v}$$

and thus $P(\lambda)$ is an eigenvalue of $P(\mathbf{A})$. For any matrix, \mathbf{A} , the sum of eigenvalues is equal to the trace of the matrix

$$\sum_{k=1}^n \lambda_k = \text{Tr}(\mathbf{A})$$

when multiplicities are taken into account. For the matrices considered in the Lemma 34.5 all eigenvalues are distinct. Thus applying this property to the matrix $P(\mathbf{A})$ gives the desired statement.

Lemma 34.6 *A Wishart distributed matrix \mathbf{W} as defined in Definition 34.2 will be a symmetric $n \times n$ matrix.*

Proof From the definition \mathbf{W} is a $p \times p$ matrix such that $\mathbf{W} = \mathbf{X}\mathbf{X}^\top$. Then

$$\mathbf{W}^\top = (\mathbf{X}\mathbf{X}^\top)^\top = (\mathbf{X}^\top)^\top \mathbf{X}^\top = \mathbf{X}\mathbf{X}^\top = \mathbf{W}$$

and thus \mathbf{W} is symmetric.

Lemma 34.7 *Suppose we have a Wishart distributed matrix \mathbf{W} with the probability density function of its eigenvalues given by*

$$\mathbb{P}(\lambda) = C_n v_n(\lambda)^m \exp\left(-\frac{\beta}{2} \sum_{k=1}^n P(\lambda_k)\right) \tag{34.23}$$

where C_n is a normalising constant, m is a positive integer, $\beta > 1$ and P is a polynomial with real coefficients. Then the vector of eigenvalues of \mathbf{W} will lie on the surface defined by

$$\sum_{k=1}^n P(\lambda_k) = \text{Tr}(P(\mathbf{W})). \tag{34.24}$$

Proof Since \mathbf{W} is symmetric by Lemma 34.6 then it will also have real eigenvalues. By Lemma 34.5

$$\sum_{k=1}^n P(\lambda_k) = \text{Tr}(P(\mathbf{W}))$$

and thus the point given by $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ will be on the surface defined by

$$\sum_{k=1}^n P(\lambda_k) = \text{Tr}(P(\mathbf{W})).$$

To find the maximum values we can use the method of Lagrange multipliers and find eigenvectors such that

$$\frac{\partial \mathbb{P}}{\partial \lambda_k} = \eta \frac{\partial}{\partial \lambda_k} \left(\text{Tr}(P(\mathbf{W})) - \sum_{k=1}^n P(\lambda_k) \right) = -\eta \frac{dP(\lambda_k)}{d\lambda_k}, \quad k = 1, \dots, n,$$

where η is some real-valued constant. Computing the left-hand side gives

$$\frac{\partial \mathbb{P}(\beta)}{\partial \lambda_k} = \mathbb{P}(\lambda) \left(-\frac{\beta}{2} \frac{dP(\lambda_k)}{d\lambda_k} + \sum_{\substack{i=1 \\ i \neq k}}^n \frac{m}{\lambda_k - \lambda_i} \right).$$

Thus the stationary points of (34.23) on the surface given by (34.24) are the solution to the equation system

$$\mathbb{P}(\lambda) \left(-\frac{\beta}{2} \frac{dP(\lambda_k)}{d\lambda_k} + \sum_{\substack{i=1 \\ i \neq k}}^n \frac{m}{\lambda_k - \lambda_i} \right) = -\eta \frac{dP(\lambda_k)}{d\lambda_k}, \quad k = 1, \dots, n.$$

If we denote the value of \mathbb{P} in a stationary point with P_s then the system above can be rewritten as

$$\sum_{\substack{i=1 \\ i \neq k}}^n \frac{1}{\lambda_k - \lambda_i} = \frac{1}{m} \left(\frac{\beta}{2} - \frac{\eta}{P_s} \right) \frac{dP(\lambda_k)}{d\lambda_k} = \rho \frac{dP(\lambda_k)}{d\lambda_k}, \quad k = 1, \dots, n. \quad (34.25)$$

The equation system described by (34.25) appears when one tries to optimize the Vandermonde determinant on a surface defined by a univariate polynomial. This problem also appears in other settings, such as finding the Fekete points on a surface [34], certain electrostatics problems [17] and D-optimal design [35]. This equation system can be rewritten as an ordinary differential equation.

Consider the polynomial

$$f(\lambda) = \prod_{i=1}^n (\lambda - \lambda_i)$$

and note that

$$\frac{1}{2} \frac{f''(\lambda_j)}{f'(\lambda_j)} = \sum_{\substack{i=1 \\ i \neq j}}^n \frac{1}{\lambda_j - \lambda_i}.$$

Thus in each of the extreme points we will have the relation

$$\left. \frac{d^2 f}{d\lambda^2} \right|_{\lambda=\lambda_j} - 2\rho \left. \frac{dP}{d\lambda} \right|_{\lambda=\lambda_j} \left. \frac{df}{d\lambda} \right|_{\lambda=\lambda_j} = 0, \quad j = 1, 2, \dots, n$$

for some $\rho \in \mathbb{R}$. Since each λ_j is a root of $f(\lambda)$ we see that the left hand side in the differential equation must be a polynomial with the same roots as $f(\lambda)$, thus we can conclude that for any $\lambda \in \mathbb{R}$

$$\frac{d^2 f}{d\lambda} - 2\rho \frac{dP}{d\lambda} \frac{df}{d\lambda} - Q(\lambda)f(\lambda) = 0 \tag{34.26}$$

where Q is a polynomial of degree $(\deg(p) - 2)$.

Consider the β ensemble described by (34.17). For this ensemble the polynomial that defines the surface that the eigenvalues will be on is $p(\lambda) = \lambda^2$. Thus by Lemma 34.7 the surface becomes a sphere with radius $\sqrt{\text{Tr}(\mathbf{W}^2)}$.

The solution to the equation system given by (34.25) on the unit sphere has been known for a long time, see [46] or [31, 35] for a more explicit description. The solution is given as the roots of a polynomial, in this case the solution can be written as the roots of the rescaled Hermite polynomials, the explicit expression for the polynomial whose roots give the maximum points is

$$\begin{aligned} f(x) &= H_n \left(\left(\frac{n-1}{2(r_1^2 - 2r_0)} \right)^{\frac{1}{2}} \frac{(x+r_1)}{2} \right) \\ &= n! \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^i}{i!} \left(\frac{n-1}{2(r_1^2 - 2r_0)} \right)^{\frac{n-2i}{2}} \frac{(x+r_1)^{n-2i}}{(n-2i)!} \end{aligned} \tag{34.27}$$

where H_n denotes the n th (physicist) Hermite polynomial [2].

The solution on the unit sphere can then be used to find the vector of eigenvalues that maximizes the probability density function $\mathbb{P}(\lambda)$ given by (34.17). Since rescaling the vector of eigenvalues affects the probability density depending on the length of the original vector in the following way

$$\mathbb{P}(c\lambda) = c^{\frac{n(n-1)m}{2}} \exp\left(\frac{\beta}{2}(1-c^2)|\lambda|^2\right) \mathbb{P}(\lambda)$$

the unit sphere solution can be rescaled so that it ends up on the appropriate sphere.

For other polynomials that define the surface that the eigenvalues lie on similar techniques, for instance for some polynomials of the form $P(\lambda) = \lambda^k$ where k is an even positive integer techniques like the ones demonstrated in [35] or [34] can be employed.

For a β ensemble the extreme points of \mathbb{P} share the properties of the extreme points of the Vandermonde determinant, for example all the extreme points will lie on the intersection of the sphere and the plane $\sum_{k=1}^n \lambda_k = 0$. What this can look like for $n = 3$ is shown in Fig. 34.1 and for $n = 4$ in Fig. 34.2.

To visualize the location of the extreme points we use a technique described in detail in [31].

It can be shown that the extreme points of $v_4(\mathbf{x})$ on the sphere all lie in the hyperplane $x_1 + x_2 + x_3 + x_4 = 0$. The intersection of this hyperplane with the unit sphere in \mathbb{R}^4 can be described as a unit sphere in \mathbb{R}^3 , under a suitable basis, and can then be easily visualized.

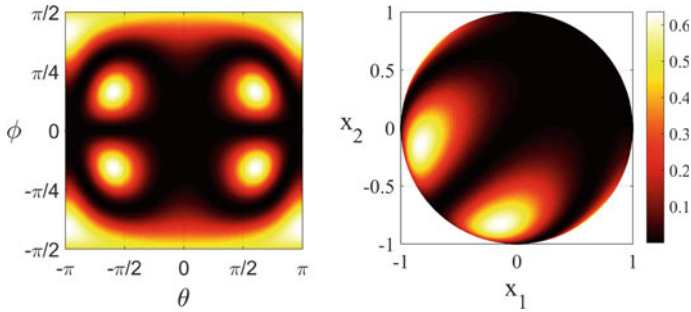


Fig. 34.1 Illustration of the expression given by (34.23) on the unit sphere in three dimensions, with parameters $n = 3, m = 2, \beta = 2$ and $C_n = 1$. Note that this expression is not correctly normalized and therefore not the exact value of the probability density distribution. On the right the value of the expression on the sphere is drawn and on the left the sphere has been parametrized in such a way that the point of the sphere given by $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ corresponds to the point (0, 0)

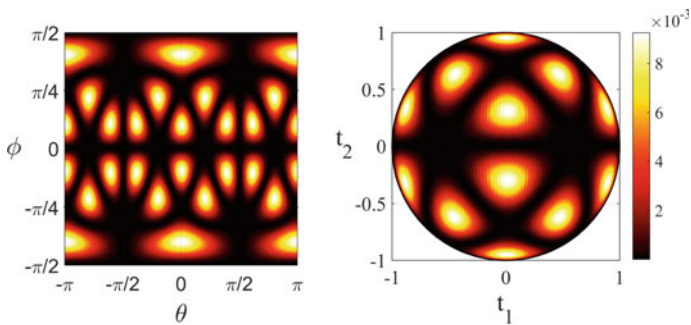


Fig. 34.2 Illustration of the expression given by (34.23) on the unit sphere in four dimensions, with parameters $n = 4, m = 2, \beta = 2$ and $C_n = 1$. Note that this expression is not correctly normalized and therefore not the exact value of the probability density distribution. In order to visualize the locations of the extreme points in four dimensions using only a two-dimensional surface the transformation given in (34.28) is used

This can be realized using the transformation

$$\mathbf{x} = \begin{pmatrix} -1 & -1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{4} & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1/\sqrt{2} \end{pmatrix} \mathbf{t} \tag{34.28}$$

where \mathbf{x} is the coordinate vector in \mathbb{R}^4 and \mathbf{t} is the corresponding coordinate vector in \mathbb{R}^3 . This will give a new sphere that can be parametrised using angles as normal.

Similar visualizations of the locations of extreme points on the unit sphere can be constructed up to $n = 7$, see [31] for further discussion.

34.6 Summary

In our study we establish that finding the extreme points for the probability distribution of the eigenvalues of a Wishart matrix can be done by finding the extreme points on a sphere with a radius related to the trace of the Wishart matrix. This close link between the Vandermonde determinant and the joint eigenvalue probability density function for β -ensembles helps to study more properties and applications of the probability density functions that occur in random matrices. As illustrated in Fig. 34.1, such results can be used to explain the distribution of charges over a unit sphere which agrees with Coloumb's theory for electrostatic charge distribution. In this case the extreme points of the probability density function of the eigenvalues happen to be the zeros of deformed Hermite polynomials given by (34.27).

Acknowledgements We acknowledge the financial support for this research by the Swedish International Development Agency, (Sida), Grant No.316, International Science Program, (ISP) in Mathematical Sciences, (IPMS). We are also grateful to the Division of Applied Mathematics, Mälardalen University for providing an excellent and inspiring environment for research education and research.

References

1. Aaronson, S., Arkhipov, A.: The computational complexity of linear optics. *Theory. Comput.* **9**, 333–342 (2013)
2. Abramowitz, M., Stegun, I.: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York (1964)
3. Anderson, G.W., Guionnet, A., Zeitouni, O.: *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics, vol. 118. Cambridge University Press (2010)
4. Anderson, T.W.: *An Introduction to Multivariate Statistical Analysis*, 3rd edn. Wiley, Publication (2003)
5. Anderson, T.W., Girshick, M.A.: Some extensions of Wishart distribution. *Ann. Math. Statist.* **15**(4), 345–357 (1944)
6. Bai, Z., Fang, Z., Liang, Y.C.: *Spectral Theory of Large Dimensional Random Matrices and Its application to Wireless Communication and Finance: Random Matrix Theory and Its Applications*. World Scientific Publishing Co., Pte., Ltd. (2014)
7. Bahcall, S.R.: Random Matrix Model for Superconductors in a Magnetic Field. *Phys. Rev. Lett.* **77**(26), 5276–5279 (1996)
8. Bleher, P.M., Its, A.R. (eds.): *Random Matrix Models and Their Applications*, MSRI Publications, vol. 40. Cambridge University Press (2001)
9. Bohigas, O., Giannoni, M.J., Schmit, S.: Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Phys. Rev. Lett.* **52**(1), 1–4 (1984)
10. Callaway, D.J.: Random matrices, fractional statistics, and the quantum Hall effect. *Phys. Rev. B* **43**(10), 8641–8643 (1991)
11. Chow, G.P.: *Analysis and Control of Dynamic Economic Systems*. Wiley, New York (1976). ISBN 0-471-15616-7
12. Cotler, J., Hunter–Jones, N., Liu, J., Yoshida, B.: Chaos, Complexity and Random Matrices. *J. High. Energy Phys.* **2017**(48) (2017)
13. del Molino, L.C.G., Luis, C., Khashayar, P., Touboul, J., Wainrib, G.: Synchronization in random balanced networks. *Phys. Rev. E* **88**(4), 042824 (2013)

14. Derrida, B.: Random–energy model: limit of a family of disordered models. *Phys. Rev. Lett.* **45**(2), 79 (1980)
15. Di Francesco, P.: 2D Quantum gravity, matrix models and graph combinatorics. In: Brezin E., Kazakov V., Serban D., Wiegmann P., Zabrodin A. (eds.), *Applications of Random Matrices in Physics*. NATO Science Series II: Mathematics, Physics and Chemistry, vol 221, 33–88, Springer, Dordrecht (2006)
16. Dumitriu, I., Edelman, A.: Matrix models for beta ensembles. *J. Math. Phys.* **43**(11), 5830–5847 (2002)
17. Dimitrov, D.K., Shapiro, B.: Electrostatic problems with a rational constraint and degenerate Lamé equations. *Potential Anal.* **52**, 645–659 (2020)
18. Edelman, A., Rao, N.R.: Random matrix theory. *Acta. Numer.* **14**, 233–297 (2005)
19. Efron, B., Morris, C.N.: Stein’s paradox in statistics. *Sci. Am.* **236**(5), 119–127 (1977)
20. Efron, B., Morris, C.N.: Multivariate empirical Bayes and estimation of covariance matrices. *Ann. Stat.* **4**(1), 22–32 (1976)
21. Franchini, F., Kravtsov, V.E.: Horizon in random matrix theory, the Hawking radiation, and flow of cold atoms. *Phys. Rev. Lett.* **103**(16), 166401 (2009)
22. Girko, V.L.: *Theory of Random Determinants*. Kluwer Academic Publishers (1990)
23. Harnad, J.: *Random Matrices, Random Processes and Integral Systems*. CRM–Series in Mathematical Physics, Springer Science and Business Media (2011)
24. James, A.T.: The distribution of latent roots of the covariance matrix. *Ann. Math. Statist.* **31**(1), 151–158 (1960)
25. Janssen, M., Pracz, K.: Correlated random band matrices: localization-delocalization transitions. *Phys. Rev. E.* **62**(6), 6278–6286 (2000)
26. Kanaka, R., Abbott, L.: Eigenvalue spectra of random matrices for neural networks. *Phys. Rev. Lett.* **97**(18), 188104 (2006)
27. Keating J (1993) The Riemann zeta-function and quantum chaos. In: *Quantum Chaos*. School of Physics Enrico Fermi, vol. CXIX, 145–185. Elsevier
28. Kemal, S.M.: *Universality in Random Matrix Models of Quantum Chromodynamics*. Doctoral Dissertation, 91191 State University of New York (1999)
29. König, W.: Orthogonal polynomial ensembles in probability theory. *Probab. Surv.* **2**, 385–447 (2005)
30. Lubinsky, D.S.: A survey of weighted polynomial approximation with exponential weights. *Surv. Approx. Theory.* **3**, 1–105 (2007)
31. Lundengård, K., Österberg, J., Silvestrov, S.: Extreme points of the Vandermonde determinant on the sphere and some limits involving the generalized Vandermonde determinant, In: Silvestrov, S., Malyarenko, A., Rančić, M. (eds.), *Algebraic Structures and Applications*, Springer Proceedings in Mathematics and Statistics, vol. 317. Springer (2020). *arXiv*, eprint [arXiv:1312.6193](https://arxiv.org/abs/1312.6193)
32. Mehta, M.L.: *Random Matrices and the Statistical Theory of Energy Levels*. Academic Press, New York, London (1967)
33. Markowitz, H.: Portfolio selection. *J. Financ.* **7**(1), 77–91 (1952)
34. Muhumuza, A.K., Lundengård, K., Österberg, J., Silvestrov, S., Mango, J.M., Kakuba, G.: The generalized Vandermonde interpolation polynomial based on divided differences. In: Skidas, C. H. (ed.), *Proceedings of the 5th Stochastic Modeling Techniques and Data Analysis International Conference with Demographics Workshop*, Chania, Crete, Greece, 2018, ISAST: International Society for the Advancement of Science and Technology, 443–456 (2018)
35. Muhumuza, A.K., Lundengård, K., Österberg, J., Silvestrov, S., Mango, J.M., Kakuba, G.: Extreme points of the Vandermonde determinant on surfaces implicitly determined by a univariate polynomial, In: Silvestrov, S., Malyarenko, A., Rančić, M. (eds.), *Algebraic Structures and Applications*, Springer Proceedings in Mathematics and Statistics, vol. 317. Springer (2020)
36. Muir, D., Mrcic-Flogel, T.: Eigenspectrum bounds for semirandom matrices with modular and spatial structure for neural networks. *Phys. Rev. E* **91**(4), 042808 (2015)
37. Muirhead, R.J.: *Aspect of Multivariate Statistical Theory*, vol. 197. Wiley (1982)

38. Novak, J.I.: Topics in Combinatorics and Random Matrix Theory. Kingston, Ontario, Canada (2009)
39. Parlett, B.N.: The Symmetric Eigenvalue Problem, vol. 20. The Society for Industrial and Applied Mathematics (SIAM) (1998)
40. Pearson, K.: On the criterion that a certain given system of deviations from the probable in the case of correlated system of variables is such that it can reasonably supposed to have arisen from random sampling. *Pil. Mg.* **50**(302), 157–175 (1900)
41. Russel, N., Chakhmakhchyan, I., O'Brien, J., Laing, A.: Direct dialling of Haar random unitary matrices. *New J. Phys.* **9**(3), 033007 (2017)
42. Rychkov, V.S., Borlenghi, S., Jaffres, H., Fert, A., Waintal, X.: Spin torque and waviness in magnetic multilayers: a bridge between Valet-Fert theory and quantum approaches. *Phys. Rev. Lett.* **103**(6), 066602 (2009)
43. Sánchez, D., Büttiker, M.: Magnetic-field asymmetry of nonlinear mesoscopic transport. *Phys. Rev. Lett.* **93**(10), 106802 (2004)
44. Sompolinsky, H., Crisanti, A., Sommers, H.: Chaos in random neural networks. *Phys. Rev. Lett.* **61**(3), 259–262 (1988)
45. Stein, C.: Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution. Stanford University Stanford, United States (1956)
46. Szegő, G.: Orthogonal Polynomials. American Mathematics Society (1939)
47. Timme, M., Wolf, F., Geisel, T.: Topological speed limits to network synchronization. *Phys. Rev. Lett.* **92**(7), 074101 (2004)
48. Tropp, J.: User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.* **12**, 389–434 (2011)
49. Turnovsky, S.: The stability properties of optimal economic policies. *Rev. Econ. Stud.* **64**(1), 136–148 (1974)
50. Turnovsky, S.: Optimal stabilization policies for stochastic linear systems: the case of correlated multiplicative and additive disturbances. *Rev. Econ. Stud.* **43**(1), 191–194 (1976)
51. Van der Vaart, A.W.: Asymptotic Statistics, vol. 3. Cambridge University Press (2000)
52. Verbaarschot, J.J., Wettig, T.: Random matrix theory and chiral symmetry in QCD. *Annu. Rev. Nucl. Part. Sci.* **50**(1), 343–410 (2000)
53. Wainrib, G., Touboul, J.: Topological and dynamical complexity of random neural networks. *Phys. Rev. Lett.* **110**(11), 118101 (2013)
54. Wigner, E.P.: Random Matrices in Physics. *SIAM Review* **9**(1), 1–23 (1967)
55. Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimension. *Ann. Math.* **62**(3), 524–540 (1955)
56. Wishart, J.: The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* **20A**(1/2), 32–52 (1928)
57. Zanon, N., Pichard, J.-L.: Random matrix theory and universal statistics for disordered quantum conductors with spin-dependent hopping. *J. Phys.* **49**(6), 907–920 (1988)
58. Ziegler, K.: Random matrix approach to light scattering on complex particles. In: The Fifth International Kharkov Symposium on Physics and Engineering of Microwaves, Millimeter, and Submillimeter Waves (IEEE Cat. No. 04EX828), Ukraine, 21–26 June 2004, vol. 1, 208–210 (2004)
59. Zumbühl, D.M., Miller, J.B., Marcus, C.M., Campman, K., Gossard, A.C.: Spin-orbit coupling, antilocalization, and parallel magnetic fields in quantum dots. *Phys. Rev. Lett.* **89**(27), 276803 (2002)