Uniform approximation of eigenvalues in Laguerre and Hermite β-ensembles by roots of orthogonal polynomials

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
holger.dette@ruhr-uni-bochum.de

Lorens A. Imhof
Department of Statistics
Bonn University
D-53113 Bonn, Germany

October 21, 2004

Abstract

We derive strong uniform approximations for the eigenvalues in general Laguerre and Hermite β-ensembles by showing that the maximal discrepancy between the suitably scaled eigenvalues and roots of orthogonal polynomials converges almost surely to zero when the dimension converges to infinity. We also provide estimates of the rate of convergence. In the special case of a normalized real Wishart matrix $W(I_n, s)/s$, where n denotes the dimension and s the degrees of freedom, the rate is $(\log n/s)^{1/4}$, if $n, s \to \infty$ with $n \leq s$, and the rate is $\sqrt{\log n/n}$, if $n, s \to \infty$ with $n \leq s \leq n + K$. In the latter case we also show the a.s. convergence of the $\lfloor nt \rfloor$ largest eigenvalue of $W(I_n, s)/s$ to the corresponding quantile of the Marčenko-Pastur law.

AMS Subject Classification: Primary 60F15, 15A52. Secondary 82B10.
Keywords and Phrases: Gaussian ensemble, random matrix, rate of convergence, Weyl's inequality, Wishart matrix.

1 Introduction

The study of random matrices has a long history in physics and statistics. Gaussian (or Hermite) ensembles arise in physics and were identified by Dyson (1962) in terms of their invariance properties, that is: Gaussian Orthogonal ensembles with real entries (GOE), Gaussian Unitary ensembles with complex entries (GUE) and Gaussian Symplectic ensembles with quaternion entries (GSE). The Wishart (or Laguerre) ensembles appear in statistics [see Muirhead (1982)] and similarly as in the Gaussian case, matrices with real, complex and quaternion entries are studied in the literature. Analytic formulas for the density of the joint distribution of the eigenvalues of such matrices were derived by

Dyson (1962) for the Hermite case and by Fisher (1939), Hsu (1939) and James (1964) for the Laguerre case, and by now it is current practice in standard random matrix models to specify the probability density of the eigenvalues without mentioning the random matrix explicitly. The numerical type of the matrix elements in these density formulas appears only as an exponent of a Vandermonde determinant, which is usually denoted by β and attains the values $1, 2, 4$, corresponding to the real, complex or quaternion case.

The Laguerre ensemble is defined by specifying the density of the joint distribution of the real eigenvalues as (up to a normalizing constant)

$$\prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^{n} \lambda_i^{a-(n-1)\frac{\beta}{2}-1} e^{-\sum_{i=1}^{n} \frac{\lambda_i}{2}} , \tag{1.1}$$

where $a > (n-1)\frac{\beta}{2} > 0$. Properties of random variables with density (1.1) have been studied by numerous authors mainly for the real ($\beta = 1$) and complex ($\beta = 2$) case [see e.g. Marčenko and Pastur (1967), Silverstein (1985), Johnstone (2001) among many others]. Because the function in (1.1) is (up to a constant) the density of the sample covariance matrix of a normally distributed sample, most asymptotic results have been transferred to the situation of a not normally distributed sample [see Bai and Yin (1988a,b, 1993) among others]. Similarly, the density corresponding to the Gaussian ensemble is proportional to

$$\prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta e^{-\sum_{i=1}^{n} \frac{\lambda_i^2}{2}} , \tag{1.2}$$

and has been studied extensively in the literature [see Mehta (1967)]. Throughout this paper we call this ensemble the Hermite-ensemble in order to emphasize the close connection to the Hermite-polynomials. The formulas (1.1) and (1.2) can obviously be extended to more general values for the exponent of the Vandermonde determinant, say $\beta > 0$, but it was unknown whether matrix models with such eigenvalue distributions exist. Recently Dumitriu and Edelmann (2002) introduced a class of random matrices such that for any $\beta > 0$ the joint eigenvalue distribution is given by the densities specified in (1.1) and (1.2). Their work was motivated by physical considerations, where the parameter β can be interpreted as inverse temperature and the cases $\beta = 0$ and $\beta = \infty$ correspond to complete independence and a frozen state, respectively.

The present paper is concerned with the asymptotic behaviour of random eigenvalues governed by the law with density (1.1) or (1.2) for arbitray $\beta > 0$. One reason for our interest in these asymptotics stems from the study of sample covariance matrices in statistics. While most work in this context deals with asymptotic properties of the empirical spectral distribution

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I\{\lambda_i \leq x\}$$

[see Marčenko and Pastur (1967), Bai (1999) or Bai, Miao and Yao (2004) among others] or the behaviour of the largest eigenvalue [see Silverstein (1985), Tracy and Widom

(2000), Johnstone (2001)], the purpose of the present paper is to study the asymptotic properties of each eigenvalue directly. Silverstein (1985) proved almost sure convergence of the (appropriately scaled) largest and smallest eigenvalue of a Wishart matrix to the boundary of the support of the Marčenko-Pastur law [for a generalization of his results to arbitrary covariance matrices see Bai and Yin (1993)], but less seems to be known about the other eigenvalues. In this paper we compare the random (scaled) eigenvalues, $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ governed by the law with density proportional to (1.1) or (1.2) with the roots $x_1 < x_2 < \ldots < x_n$ of appropriately scaled Laguerre or Hermite polynomials, respectively, where the parameter $\beta > 0$ is arbitrary (i.e. it is not necessarily equal to $1, 2$ or 4). It is well known that there is a close connection between random matrix theory and the theory of orthogonal polynomials [see e.g. Deift (1998)]. We derive explicit bounds for the probability

$$P\Big\{\max_{i=1}^{n} |\lambda_i - x_i| > \epsilon\Big\} \tag{1.3}$$

and establish the almost sure convergence of $\max_{i=1}^{n} |\lambda_i - x_i|$ with a nearly optimal rate. Our approach heavily relies on specific matrix models, which were recently introduced by Dumitriu and Edelman (2002) and yield a joint eigenvalue distribution of the form (1.1) or (1.2) for any $\beta > 0$. Our bounds of the probability (1.3) also allow to derive convergence results with explicit rates for the eigenvalues of random matrices of a fixed dimension as $\beta \to \infty$. The rates thus obtained turn out to be optimal again.
Section 2 deals with the general Laguerre ensemble, while we derive in Section 3 strong uniform approximations of the eigenvalues of a Wishart matrix $W(I_n, s)/s$ by roots of the Laguerre polynomial $L_n^{(s-n)}(sx)$ with rate $(\log n/s)^{1/4}$, if $n, s \to \infty$ with $n \le s$, and with rate $\sqrt{\log n/n}$, if $n, s \to \infty$ with $n \le s \le n + K$, where n denotes the dimension and s the degrees of freedom tending to infinity. In the latter case we also show the a.s. convergence of the $\lfloor nt \rfloor$th largest eigenvalue of the matrix $W(I_n, s)/s$ to the corresponding quantile of the Marčenko and Pastur law with rate $\sqrt{\log n/n}$. This generalizes a result of Silverstein (1985), who considered only the smallest and largest eigenvalue and did not derive the rate of convergence. Finally, in Section 4 we turn to the general β-Hermite ensemble, while some technical details are presented in the appendix.

2 Laguerre Ensembles

Recall the definition (1.1) of the β-Laguerre ensemble, where the parameter β varies in the interval $(0, \infty)$. We first study for every fixed dimension $n \ge 2$ the maximal distance between the random eigenvalues $\lambda_1, \ldots, \lambda_n$ corresponding to a β-Laguerre ensemble (scaled by $\frac{1}{2a}$) and roots of a suitably scaled Laguerre polynomial. To this end, we make use of the following random matrix model, which was recently introduced by Dumitriu and Edelman (2002). Let $a, \beta \in \mathbb{R}$, where $\beta > 0$ and

$$a > \frac{\beta}{2}(n-1) \tag{2.1}$$

and define $X_{2a}, X_{2a-\beta}, \ldots, X_{2a-(n-1)\beta}, Y_\beta, Y_{2\beta}, \ldots, Y_{(n-1)\beta}$ as independent random variables with $X_r^2 \sim \chi^2(r)$, $X_r \ge 0$ and $Y_r^2 \sim \chi^2(r)$, $Y_r \ge 0$, where $\chi^2(r)$ denotes a χ-square distribution with r degrees of freedom. The scaled Laguerre matrix $M = M(n, a, \beta)$ is defined by

$$M = M(n, a, \beta) = \frac{1}{2a} BB^T, \tag{2.2}$$

with

$$B = \begin{bmatrix} X_{2a} & & & \\ Y_{(n-1)\beta} & X_{2a-\beta} & & \\ & \ddots & \ddots & \\ & & Y_\beta & X_{2a-(n-1)\beta} \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Note that M is tridiagonal and that it was shown by Dumitriu and Edelman (2002) that the joint density of the eigenvalues of the matrix BB^T is proportional to the function defined by (1.1). In the following we denote by $\lambda_1 \le \cdots \le \lambda_n$ the (ordered) eigenvalues of the matrix M. For $\alpha > -1$, let $L_n^{(\alpha)}(x)$ denote the nth Laguerre polynomial orthogonal with respect to the weight $x^\alpha e^{-x}$ on the interval $(0, \infty)$ and define $x_1 < \cdots < x_n$ as the (ordered) zeros of the scaled Laguerre polynomial $L_n^{((2a/\beta)-n)}(2ax/\beta)$. Our first result gives an estimate for the probability that the maximum difference between the eigenvalues of the random matrix and the corresponding roots of the orthogonal polynomial exceeds a certain bound, say $\epsilon > 0$.

Theorem 2.1. *Let $x_1 < \cdots < x_n$ denote the zeros of the scaled Laguerre polynomial $L_n^{((2a/\beta)-n)}(2ax/\beta)$ and $\lambda_1 \le \cdots \le \lambda_n$ the eigenvalues of the matrix M defined in (2.2). Then we have for any $0 < \epsilon < 1$*

$$P\Big\{\max_{1\le j\le n} |\lambda_j - x_j| > \epsilon\Big\} \le 4n\Big[\Big(1 + \frac{\epsilon^2}{25}\Big)\exp\Big(-\frac{\epsilon^2}{25}\Big)\Big]^a.$$

Proof. Note first that the entries of the random tridiagonal matrix $M = (M_{ij})_{ij=1}^n$ are given by

$$\begin{aligned} M_{11} &= \frac{1}{2a} X_{2a}^2, \\ M_{ii} &= \frac{1}{2a}\left[X_{2a-(i-1)\beta}^2 + Y_{(n+1-i)\beta}^2\right], \qquad & i = 2, \ldots, n, \\ M_{i,i+1} = M_{i+1,i} &= \frac{1}{2a} X_{2a-(i-1)\beta} Y_{(n-i)\beta}, \qquad & i = 1, \ldots, n-1, \end{aligned}$$

and consider the deterministic symmetric tridiagonal matrix $C_n = (c_{ij})_{i,j=1}^n$ with entries

$$\begin{aligned} c_{ii} &= \frac{2a + (n + 2 - 2i)\beta}{2a}, \\ c_{i,i+1} &= c_{i+1,i} = \frac{1}{2a}\sqrt{[2a - (i-1)\beta]\,(n-i)\beta}. \end{aligned}$$

Let C_{n-1} denote the matrix obtained from C_n by deleting the first column and first row. Using the recurrence relation for Laguerre polynomials [see Szegö (1975), (5.1.10), page

101, or Chihara (1978), page 220] one may verify that

$$\det(xI_n - C_n) = \left(\frac{\beta}{2a}\right)^n \hat{L}_n^{((2a/\beta)-n+1)}(2ax/\beta),$$
$$\det(xI_{n-1} - C_{n-1}) = \left(\frac{\beta}{2a}\right)^{n-1} \hat{L}_{n-1}^{((2a/\beta)-n+1)}(2ax/\beta),$$

where I_k is the $k \times k$ identity matrix and $\hat{L}_k^{(\alpha)}(x) = (-1)^k k! L_k^{(\alpha)}(x)$ is the kth monic Laguerre polynomial. Let $\mathbf{e}_1 = (1, 0, \ldots, 0)^T$ be the first unit vector in $\mathbb{R}^n$ and define $D := C_n - [n\beta/(2a)]\mathbf{e}_1\mathbf{e}_1^T$. Then

$$\begin{aligned}
\det(xI_n - D) &= \det(xI_n - C_n) + \frac{n\beta}{2a}\det(xI_{n-1} - C_{n-1}) \\
&= \left(\frac{-\beta}{2a}\right)^n n! \left\{ L_n^{(2a/\beta-n+1)}(2ax/\beta) - L_{n-1}^{(2a/\beta-n+1)}(2ax/\beta) \right\} \\
&= \left(\frac{-\beta}{2a}\right)^n n! L_n^{(2a/\beta-n)}(2ax/\beta)
\end{aligned}$$

[see Szegö (1975), (5.1.13), page 102], which proves that the points $x_1, \ldots, x_n$ are also the eigenvalues of the matrix D.

From Weyl's inequalities [Horn and Johnson (1985), Theorem 4.3.1] we therefore obtain

$$\max_{1 \le j \le n} |\lambda_j - x_j| \le \rho\left(M - D\right), \tag{2.3}$$

where

$$\rho(A) = \max\{|\mu| : \mu \text{ is an eigenvalue of } A\}$$

denotes the spectral radius of a matrix $A \in \mathbb{R}^{n \times n}$. Let

$$\|A\|_\infty := \max_{1 \le i \le n} \sum_{j=1}^n |a_{ij}|.$$

According to Theorem 5.6.9 in Horn and Johnson (1985),

$$\rho(A) \le \|A\|_\infty$$

for all $A \in \mathbb{R}^{n \times n}$, and it therefore follows from (2.3) that

$$\max_{1 \le j \le n} |\lambda_j - x_j| \le \|M - D\|_\infty. \tag{2.4}$$

With the notation

$$Z_n := \max\left\{ \max_{0 \le i \le n-1} \frac{|X_{2a-i\beta}^2 - (2a - i\beta)|}{2a}, \max_{1 \le i \le n-1} \frac{|Y_{i\beta}^2 - i\beta|}{2a} \right\},$$

we obtain for the elements of the first row of the matrix $M - D$ that

$$\begin{aligned}
\sum_{j=1}^n \left|\{M - D\}_{1j}\right| &= \left|M_{11} - (c_{11} - \frac{n\beta}{2a})\right| + |M_{12} - c_{12}| \\
&= \frac{1}{2a}\left|X_{2a}^2 - 2a\right| + \frac{1}{2a}\left|X_{2a}Y_{(n-1)\beta} - \sqrt{2a}\sqrt{(n-1)\beta}\right| \\
&\le 2Z_n + \left\{1 + \sqrt{\frac{(n-1)\beta}{2a}}\right\}\sqrt{Z_n} \le 2Z_n + 2\sqrt{Z_n}\ ,
\end{aligned}$$

where we used assumption (2.1) and the inequality

$$|xy - \underline{x}\,\underline{y}| \leq |x^2 - \underline{x}^2|^{1/2}|y^2 - \underline{y}^2|^{1/2} + |\underline{x}||y^2 - \underline{y}^2|^{1/2} + |\underline{y}||x^2 - \underline{x}^2|^{1/2}$$

for $x, y, \underline{x}, \underline{y} \geq 0$, see Silverstein (1985). Similarly, it follows for the elements in the rows $2, \ldots, n$ of the matrix $M - D$ that

$$\sum_{j=1}^{n} \left| \{M - D\}_{ij} \right| \leq 4Z_n + 4\sqrt{Z_n}, \qquad i = 2, \ldots, n.$$

Hence $\|M - D\|_\infty \leq 4Z_n + 4\sqrt{Z_n}$, and therefore we obtain from (2.4) that

$$\max_{1 \leq j \leq n} |\lambda_j - x_j| \leq 4\left(Z_n + \sqrt{Z_n}\right).$$

Consequently, for $0 < \epsilon < 1$,

$$P\Big\{\max_{1 \leq j \leq n} |\lambda_j - x_j| > \epsilon\Big\} \leq P\Big\{4Z_n + 4\sqrt{Z_n} > \epsilon\Big\} \leq P\Big\{5\sqrt{Z_n} > \epsilon\Big\}, \qquad (2.5)$$

since $4z \leq \sqrt{z}$ for $0 \leq z \leq \frac{1}{16}$, and $5\sqrt{z} > \epsilon$ for $z > \frac{1}{16}$. It is clear from the definition of random variable Z_n that

$$P\Big\{Z_n > \frac{\epsilon^2}{5^2}\Big\} \leq \sum_{i=0}^{n-1} P\Big\{\frac{|X^2_{2a-i\beta} - (2a - i\beta)|}{2a} > \frac{\epsilon^2}{5^2}\Big\} + \sum_{i=1}^{n-1} P\Big\{\frac{|Y^2_{i\beta} - i\beta|}{2a} > \frac{\epsilon^2}{5^2}\Big\}.$$

For $i = 1, \ldots, n-1$ we obtain $i\beta \leq 2a$, by assumption (2.1), and therefore it follows from Lemma A.1 (v) in the appendix that

$$P\Big\{\frac{|Y^2_{i\beta} - i\beta|}{2a} > \frac{\epsilon^2}{5^2}\Big\} \leq 2\Big(1 + \frac{2a\epsilon^2}{25 i\beta}\Big)^{i\beta/2} \exp\left(-\frac{a\epsilon^2}{25}\right) \leq 2\Big(1 + \frac{\epsilon^2}{25}\Big)^{a} \exp\left(-\frac{a\epsilon^2}{25}\right),$$

where the last inequality uses the fact that the function $(1 + c/x)^x$ is increasing for $x > 0$ $(c > 0)$. By a similar argument we have for $i = 0, \ldots, n-1$,

$$P\Big\{\frac{|X^2_{2a-i\beta} - (2a - i\beta)|}{2a} > \frac{\epsilon^2}{25}\Big\} \leq 2\Big(1 + \frac{\epsilon^2}{25}\Big)^{a} \exp\left(-\frac{a\epsilon^2}{25}\right),$$

which gives

$$P\Big\{Z_n > \frac{\epsilon^2}{25}\Big\} \leq 2(2n-1)\Big[\Big(1 + \frac{\epsilon^2}{25}\Big)\exp\left(-\frac{\epsilon^2}{25}\right)\Big]^{a}.$$

Combining this inequality with (2.5) yields the assertion of the theorem. □

Note that the bound in Theorem 2.1 depends on the parameter β only through the inequality (2.1). Moreover, if $n \to \infty$ we obtain by assumption (2.1) $n/a = \mathcal{O}(1)$ and $P\{\max_{1 \leq j \leq n} |\lambda_j - x_j| > \epsilon\}$ converges to zero with an exponential rate. The next result uses this fact and gives a strong limit theorem for the maximum of the absolute differences between the eigenvalues of general Laguerre ensembles of size n and roots of Laguerre polynomials when $n \to \infty$, where we also allow the parameters a and β to depend on n.

Theorem 2.2. *Let (a_n) and (β_n) be two sequences of parameters such that for every n, $a_n > \beta_n(n-1)/2 > 0$, and let $x_1^{(n)} < \cdots < x_n^{(n)}$ denote the zeros of the scaled Laguerre polynomial $L_n^{((2a_n/\beta_n)-n)}(2a_n x/\beta_n)$. If*

$$\lim_{n\to\infty} \frac{a_n}{\log n} = \infty \ , \tag{2.6}$$

then the eigenvalues $\lambda_1^{(n)} \le \cdots \le \lambda_n^{(n)}$ of the scaled Laguerre matrix $M(n, a_n, \beta_n)$ defined in (2.2) satisfy

$$\max_{1\le j\le n} \left|\lambda_j^{(n)} - x_j^{(n)}\right| \le \left(\frac{\log n}{a_n}\right)^{1/4} S \qquad \text{for all } n \ge 2,$$

where S denotes an a.s. finite random variable. In particular, if

$$\liminf_{n\to\infty} \frac{a_n}{n} > 0, \tag{2.7}$$

then there exists an a.s. finite random variable S' such that

$$\max_{1\le j\le n} \left|\lambda_j^{(n)} - x_j^{(n)}\right| \le \left(\frac{\log n}{n}\right)^{1/4} S' \qquad \text{for all } n \ge 2.$$

Proof. For $n \ge 2$, set

$$R_n = \Big(\frac{a_n}{\log n}\Big)^{1/4} \max_{1\le j\le n} \big|\lambda_j^{(n)} - x_j^{(n)}\big|.$$

We have to show that $\sup_n R_n$ is a.s. finite. To this end we first show that if (ϕ_n) is any non-random sequence of positive numbers with $\phi_n \to \infty$, then $R_n/\phi_n \to 0$ a.s. Fix such a sequence (ϕ_n), fix $\epsilon > 0$ and define

$$\epsilon_n = \epsilon \min\Big\{\phi_n, \Big(\frac{a_n}{\log n}\Big)^{1/8}\Big\} \Big(\frac{\log n}{a_n}\Big)^{1/4}, \qquad n \ge 2.$$

By (2.6), $\epsilon_n \to 0$. In particular, for n sufficiently large, $\epsilon_n < 1$, and so, by Theorem 2.1,

$$P\Big\{\frac{R_n}{\phi_n} > \epsilon\Big\} \le P\Big\{\max_{1\le j\le n} \big|\lambda_j^{(n)} - x_j^{(n)}\big| > \epsilon_n\Big\} \le 4n\Big[\Big(1 + \frac{\epsilon_n^2}{5^2}\Big)\exp\Big(-\frac{\epsilon_n^2}{5^2}\Big)\Big]^{a_n}.$$

The function $\psi(x) = \log(1+x) - x + x^2/4$ is increasing on the interval $(-1, 0]$ and decreasing on $[0, 1]$, so that $\psi(x) \le \psi(0) = 0$ for $x \in (-1, 1]$. Thus

$$(1+x)e^{-x} \le e^{-x^2/4}, \qquad x \le 1 \ , \tag{2.8}$$

and it follows that

$$P\Big\{\frac{R_n}{\phi_n}>\epsilon\Big\}\le 4n\exp\Big(-\frac{\epsilon_n^4 a_n}{4\cdot 5^4}\Big)=4n^{1-\epsilon^4\min\{\phi_n^4,(a_n/\log n)^{1/2}\}/2500}.$$

Using condition (2.6) again, we obtain that

$$\sum_{n=2}^{\infty}P\Big\{\frac{R_n}{\phi_n}>\epsilon\Big\}<\infty.$$

Hence, by the lemma of Borel and Cantelli it follows that $R_n/\phi_n\to 0$ a.s.

To complete the proof define

$$S_n=\max_{1\le k\le n}R_k$$

and

$$S=\sup_{n\ge 2}S_n=\sup_{n\ge 2}R_n.$$

Assume that $\delta:=P\{S=\infty\}>0$. Define a sequence (ϕ_n) by

$$\phi_n=\sup\Big\{\phi\ge 0: P\{S_n\ge\phi\}\ge\frac{\delta}{2}\Big\}.$$

As (S_n) is increasing, so is (ϕ_n). Moreover, the sequence (ϕ_n) is unbounded because the assumption $\phi_n\le\Phi$ for all n and some constant $\Phi\in\mathbb{R}$ would imply

$$\delta\le P\Big(\bigcup_{n=2}^{\infty}\{S_n\ge\Phi+1\}\Big)\le\limsup_{n\to\infty}P\{S_n\ge\phi_n+1\}\le\frac{\delta}{2};$$

which yields a contradiction. Consequently $\phi_n\to\infty$, and it follows by the first part of the proof that $R_n/\phi_n\to 0$ a.s. However, this also implies that

$$S_n/\phi_n\to 0\quad\text{a.s.}$$

To see this fix ω with $R_n(\omega)/\phi_n\to 0$. Then there exists a sequence of indices, say $k_1(\omega)\le k_2(\omega)\le\dots$, such that for every n, $k_n(\omega)\le n$ and $R_{k_n(\omega)}(\omega)=S_n(\omega)$. If the sequence $(k_n(\omega))$ is bounded, it is eventually constant, and $S_n(\omega)/\phi_n=R_{k_n(\omega)}(\omega)/\phi_n\to 0$. Otherwise,

$$\frac{S_n(\omega)}{\phi_n}=\frac{R_{k_n(\omega)}(\omega)}{\phi_{k_n(\omega)}}\frac{\phi_{k_n(\omega)}}{\phi_n}\le\frac{R_{k_n(\omega)}(\omega)}{\phi_{k_n(\omega)}}\to 0.$$

Consequently, $S_n/\phi_n\to 0$ a.s. On the other hand, in view of the definition of the sequence (ϕ_n), S_n/ϕ_n does not converge to 0 in probability. This contradiction shows that $\delta=0$, that is, $S<\infty$ a.s., and completes the proof. □

Remark 2.3. Dumitriu and Edelman (2004) describe a physical model where the density of the eigenvalues is given by (1.1), the parameter $\beta>0$ is interpreted as inverse temperature; an increase in temperature (i.e. a smaller value for the parameter β) yields a larger degree of randomness of the eigenvalues. In contrast to Theorem 2.2

these authors considered the case where the dimension is fixed and the parameters β and $a = a_\beta$ converge to infinity, such that

$$\lim_{\beta\to\infty} \frac{a_\beta}{\beta} = \frac{1}{2}(n - 1 + \gamma)$$

for some $\gamma > 0$. They proved convergence in probability of the jth eigenvalue $\bar{\lambda}_j$ of the matrix M/β to the jth root $\bar{x}_j$ of the Laguerre polynomial $L_n^{(\gamma-1)}(x)$. This statement can also easily be obtained from Theorem 2.1, which additionally shows that the corresponding probabilities decrease exponentially.
In Theorem 2.2 we consider the case where the dimension and the parameters may vary. The general assumption that $2a_n/\beta_n > n-1$ implies that if the sequence of temperatures (β_n^{-1}) is bounded, then condition (2.7) in Theorem 2.2 is satisfied. In light of the above model, it is not surprising that assumptions of this type on the sequences (a_n) or (β_n) appear in a strong limit theorem. In particular, condition (2.6) is not very restrictive.

The next aim is to improve the rate of convergence established in Theorem 2.2 under a certain restriction on the parameters a_n and β_n. For this we first prove an extension of Theorem 2.1.

Theorem 2.4. *Let $K \geq 0$ and $\epsilon \in (0,1)$, define $x_1 < \cdots < x_n$ as the zeros of the scaled Laguerre polynomial $L_n^{((2a/\beta)-n)}(2ax/\beta)$ and let $\lambda_1 \leq \cdots \leq \lambda_n$ denote the eigenvalues of the matrix M defined in (2.2). If*

$$n + K \geq \frac{2a}{\beta} \geq n - 1 + \frac{1}{\beta} \tag{2.9}$$

and

$$\frac{a\epsilon}{\beta} \geq 2(K+1), \tag{2.10}$$

then we have

$$P\Big\{\max_{1\leq j\leq n} |\lambda_j - x_j| > \epsilon\Big\} \leq 6n\Big\{\Big(1 + \frac{\epsilon}{4\sqrt{K+2}}\Big)\exp\Big(-\frac{\epsilon}{4\sqrt{K+2}}\Big)\Big\}^a$$
$$+ 2n\exp\Big\{\frac{(K+1)^2\beta}{8}\Big\}\Big(1 - \frac{\epsilon}{4}\Big)^{(n-1)\beta}\exp\Big(\frac{a\epsilon}{2}\Big).$$

Proof. Define the matrix D as in the proof of Theorem 2.1, then it follows from (2.4) that

$$\max_{1\leq j\leq n} |\lambda_j - x_j| \leq \|M - D\|_\infty.$$

With the notation

$$\begin{aligned}
Z_n^{(1)} &= \max_{0\leq i\leq n-1} \frac{|X_{2a-i\beta}^2 - (2a - i\beta)|}{2a}, \\
Z_n^{(2)} &= \max_{1\leq i\leq n-1} \frac{|Y_{i\beta}^2 - i\beta|}{2a}, \\
Z_n^{(3)} &= \max_{1\leq i\leq n-1} \frac{|X_{2a-(i-1)\beta}Y_{(n-i)\beta} - \sqrt{[2a-(i-1)\beta](n-i)\beta}|}{2a}
\end{aligned}$$

we have by a similar argument as given in the proof of Theorem 2.1

$$\|M-D\|_\infty \le 4\max\{Z_n^{(1)}, Z_n^{(2)}, Z_n^{(3)}\},$$

which implies

$$\max_{1\le j\le n} |\lambda_j - x_j| \le 4\max\{Z_n^{(1)}, Z_n^{(2)}, Z_n^{(3)}\}.$$

Therefore,

$$P\Big\{\max_{1\le j\le n} |\lambda_j - x_j| > \epsilon\Big\} \le P\left\{\max\{Z_n^{(1)}, Z_n^{(2)}\} \ge \frac{\epsilon}{4}\right\} + P\left\{Z_n^{(3)} \ge \frac{\epsilon}{4}\right\}. \qquad (2.11)$$

In what follows, we will use repeatedly that for every fixed $c > 0$ the function $[1+(c/x)]^x$ is increasing in $x > 0$ and that the function $[1 + (x/c)]^c e^{-x}$ is decreasing in $x \ge 0$. If $U_s^2 \sim \chi^2(s)$ and $s \le 2a$, then it follows from Lemma A.1 (v) in the appendix that

$$\begin{aligned} P\Big\{\frac{|U_s^2 - s|}{2a} \ge \frac{\epsilon}{4}\Big\} &\le 2\left(1+\frac{a\epsilon}{2s}\right)^{s/2} \exp\left(-\frac{a\epsilon}{4}\right) \le 2\left(1+\frac{\epsilon}{4}\right)^a \exp\left(-\frac{a\epsilon}{4}\right) \\ &\le 2\Big\{\Big(1+\frac{\epsilon}{4\sqrt{K+2}}\Big)\exp\Big(-\frac{\epsilon}{4\sqrt{K+2}}\Big)\Big\}^a =: p_1. \end{aligned}$$

Observing that for $i = 1, \ldots, n-1$, $i\beta \le 2a$, we therefore obtain

$$P\left\{\max\{Z_n^{(1)}, Z_n^{(2)}\} \ge \frac{\epsilon}{4}\right\} \le (2n-1)p_1. \qquad (2.12)$$

To determine an upper bound of the remaining probability $P\{Z_n^{(3)} \ge \epsilon/4\}$ we fix $i \in \{1, \ldots, n-1\}$ and write $q = 2a - (i-1)\beta$, $r = (n-i)\beta$. With this notation it follows that $q \le 2a$, $r \le 2a$, $q/r \ge 1/(K+2)$ and $r/q \ge 1/(K+2)$ by assumption (2.9). Using these inequalities and Lemma A.2 a) in the appendix we obtain that

$$\begin{aligned} P\Big\{\frac{X_qY_r - \sqrt{qr}}{2a} \ge \frac{\epsilon}{4}\Big\} &\le \Big(1+\frac{a\epsilon}{2q}\sqrt{\frac{q}{r}}\Big)^{q/2}\exp\Big(-\frac{a\epsilon}{4}\sqrt{\frac{q}{r}}\Big) \\ &\qquad + \Big(1+\frac{a\epsilon}{2r}\sqrt{\frac{r}{q}}\Big)^{r/2}\exp\Big(-\frac{a\epsilon}{4}\sqrt{\frac{r}{q}}\Big) \\ &\le \Big(1+\frac{\epsilon}{4}\sqrt{\frac{q}{r}}\Big)^{a}\exp\Big(-\frac{a\epsilon}{4}\sqrt{\frac{q}{r}}\Big) + \Big(1+\frac{\epsilon}{4}\sqrt{\frac{r}{q}}\Big)^{a}\exp\Big(-\frac{a\epsilon}{4}\sqrt{\frac{r}{q}}\Big) \\ &\le 2\Big\{\Big(1+\frac{\epsilon}{4\sqrt{K+2}}\Big)\exp\Big(-\frac{\epsilon}{4\sqrt{K+2}}\Big)\Big\}^a = p_1. \end{aligned}$$

Suppose for the moment that $a\epsilon/2 \le \sqrt{qr}$. Observing that $q \ge r+1$ (in view of assumption (2.9)) we have by Lemma A.2 b),

$$P\Big\{\frac{\sqrt{qr} - X_qY_r}{2a} \ge \frac{\epsilon}{4}\Big\} \le \Big(1-\frac{a\epsilon}{2\sqrt{qr}}\Big)^r \Big[\frac{1}{\sqrt{2}}\exp\Big\{\frac{1}{2}\Big(\sqrt{q}-\sqrt{r}\Big)^2\Big\}+1\Big]\exp\Big(\frac{a\epsilon}{2}\Big). \quad (2.13)$$

It follows from the mean-value theorem and assumption (2.9) that

$$(\sqrt{q}-\sqrt{r})^2 \le \frac{(q-r)^2}{4r} \le \frac{\beta}{4}\Big\{\frac{2a}{\beta} - (n-1)\Big\}^2 \le \frac{\beta}{4}(K+1)^2,$$

and therefore the term in brackets on the right-hand side of (2.13) is bounded by the expression

$$2\exp\left\{(K+1)^2\beta/8\right\}.$$

With the notation $c_1 = q-r = 2a-(n-1)\beta$ and $c_2 = a\epsilon/2$ we obtain from the assumptions (2.9) and (2.10), $c_1 \le (K+1)\beta \le c_2$. Moreover, $r \le (n-1)\beta \le 2a$, and it therefore follows from Lemma A.3 in the appendix that

$$\begin{aligned}\Big(1-\frac{a\epsilon}{2\sqrt{qr}}\Big)^r = \Big(1-\frac{c_2}{\sqrt{(r+c_1)r}}\Big)^r &\le \Big(1-\frac{c_2}{\sqrt{2a(n-1)\beta}}\Big)^{(n-1)\beta}\\ &\le \Big(1-\frac{c_2}{2a}\Big)^{(n-1)\beta} = \Big(1-\frac{\epsilon}{4}\Big)^{(n-1)\beta}.\end{aligned}$$

Hence, we have from (2.13),

$$P\Big\{\frac{\sqrt{qr}-X_qY_r}{2a} \ge \frac{\epsilon}{4}\Big\} \le 2\exp\Big\{\frac{(K+1)^2\beta}{8}\Big\}\Big(1-\frac{\epsilon}{4}\Big)^{(n-1)\beta}\exp\Big(\frac{a\epsilon}{2}\Big) =: p_2\ ,$$

in the case $a\epsilon/2 \le \sqrt{qr}$. Because this inequality is trivially true if $a\epsilon/2 > \sqrt{qr}$ we obtain

$$P\left\{Z_n^{(3)} \ge \epsilon/4\right\} \le (n-1)(p_1+p_2),$$

and the assertion of Theorem 2.4 follows from (2.11) and (2.12). □

Theorem 2.5. *Let (a_n), (β_n), $x_j^{(n)}$ and $\lambda_j^{(n)}$ be defined as in Theorem 2.2. Suppose that for some $K>0$,*

$$n+K \ge \frac{2a_n}{\beta_n} \ge (n-1)+\frac{1}{\beta_n} \qquad \textit{for every } n \ge 2. \tag{2.14}$$

Then there exists an a.s. finite variable S such that

$$\max_{1\le j\le n}\big|\lambda_j^{(n)} - x_j^{(n)}\big| \le \Big(\frac{\log n}{n}\Big)^{1/2} S \qquad \textit{for every } n \ge 2.$$

Proof. Let $(\phi_n)_{n=2}^{\infty}$ be an arbitrary sequence of positive numbers with $\phi_n \to \infty$, $\epsilon > 0$ and define

$$\begin{aligned}R_n &= \Big(\frac{n}{\log n}\Big)^{1/2}\max_{1\le j\le n}|\lambda_j^{(n)} - x_j^{(n)}|,\\ \epsilon_n &= \epsilon\min\{\phi_n, n^{1/4}\}\Big(\frac{\log n}{n}\Big)^{1/2}.\end{aligned}$$

For n sufficiently large, $\epsilon_n < 1$ and, by the second inequality in (2.14), $a_n\epsilon_n/\beta_n \ge 2(K+1)$. Hence, by Theorem 2.4 and the first inequality in (2.14),

$$\begin{aligned}P\Big\{\frac{R_n}{\phi_n} > \epsilon\Big\} &\le 6n\Big\{\Big(1+\frac{\epsilon_n}{4\sqrt{K+2}}\Big)\exp\Big(-\frac{\epsilon_n}{4\sqrt{K+2}}\Big)\Big\}^{a_n}\\ &+ 2n\exp\Big\{\frac{(K+1)^2\beta_n}{8}+\frac{(K+1)\beta_n\epsilon_n}{4}\Big\}\Big\{\Big(1-\frac{\epsilon_n}{4}\Big)\exp\Big(\frac{\epsilon_n}{4}\Big)\Big\}^{(n-1)\beta_n}\\ &=: 6c_1(n)+2c_2(n),\end{aligned}$$

where the last equality defines the functions c_1 and c_2.
From assumption (2.14) it is easy to see that

$$\text{(i) } \beta_n \geq \frac{1}{K+1}, \qquad \text{(ii) } a_n \geq \frac{n}{4(K+1)}, \tag{2.15}$$

and the inequalities (2.8) and (2.15) (ii) give

$$c_1(n) \leq n \exp\Big\{-\frac{a_n \epsilon_n^2}{4^3(K+2)}\Big\} \leq n \exp\Big\{-\frac{n\epsilon_n^2}{4^4(K+1)(K+2)}\Big\} \;=\; \mathcal{O}(n^{-2}).$$

By (2.8) and (2.15) (i) we therefore obtain

$$\begin{aligned} c_2(n) &\leq n \exp\left[\left\{\frac{(K+1)^2}{8} + \frac{(K+1)\epsilon_n}{4} - \frac{(n-1)\epsilon_n^2}{4^3}\right\}\beta_n\right] \\ &\leq n \exp\left[\frac{(K+1)}{8} + \frac{\epsilon_n}{4} - \frac{(n-1)\epsilon_n^2}{4^3(K+1)}\right], \end{aligned}$$

provided that n is so large that the term in braces is negative. It now also follows that $c_2(n) = \mathcal{O}(n^{-2})$, which implies

$$\sum_{n=2}^{\infty} P\Big\{\frac{R_n}{\phi_n} > \epsilon\Big\} < \infty.$$

The lemma of Borel and Cantelli yields that $R_n/\phi_n \to 0$ a.s. for any sequence (ϕ_n) satisfying $\phi_n \to \infty$. Finally, the assertion of Theorem 2.5 is obtained by the same argument as presented in the second part of the proof of Theorem 2.2. □

3 A brief discussion of Wishart matrices

We now present a brief discussion of the corresponding limit theorems in the important special case of real Wishart ensembles. To be precise consider for integers n, s with $n \leq s$ an $n \times s$ matrix V_s whose entries are i.i.d. $N(0,1)$ random variables and define

$$\frac{1}{s} V_s V_s^T$$

as the sample covariance matrix and note that this is a scaled Wishart matrix. Note also that the first part of the following result does not require that s/n converges.

Theorem 3.1. *Let $\lambda_1 \leq \cdots \leq \lambda_n$ denote the eigenvalues of the scaled Wishart matrix $\frac{1}{s} V_s V_s^T$ and denote by $x_1 \leq \cdots \leq x_n$ the zeros of the Laguerre polynomial $L_n^{(s-n)}(sx)$.*

(i) If $n \to \infty$ and $s = s(n) \to \infty$ such that for every n, $n \leq s(n)$, then

$$\max_{1 \leq j \leq n} |\lambda_j - x_j| = \mathcal{O}([\log n / s(n)]^{\frac{1}{4}}) \qquad a.s.$$

(ii) If $n, s(n) \to \infty$ such that for every n, $n \le s(n) \le n + K$, then

$$\max_{1 \le j \le n} |\lambda_j - x_j| = \mathcal{O}(\sqrt{\log n/n}) \qquad a.s. \tag{3.1}$$

Proof. The result follows from Theorem 2.2 and Theorem 2.5 with $\beta_n = 1$ and $a_n = s(n)/2$. $\square$

Remark 3.2. Soshnikov (2002) considered a general sample covariance matrix without the assumption of normally distributed errors. If $s/n \to \gamma > 0$ he showed for the largest eigenvalue λ_n of the sample covariance matrix the estimate

$$\lambda_n \le \frac{(\sqrt{n} + \sqrt{s})^2}{s} + \mathcal{O}\Big(\frac{\log n}{\sqrt{n}}\Big) \;\; a.s.$$

In the case $n \le s \le n + K$ considered in Theorem 3.1 his estimate reduces to

$$\lambda_n \le 4 + \mathcal{O}\Big(\frac{\log n}{\sqrt{n}}\Big) \;\; a.s.$$

In the Wishart case Theorem 3.1 yields a slightly better estimate, that is

$$\lambda_n \le 4 + \mathcal{O}\Big(\sqrt{\frac{\log n}{n}}\Big) \;\; a.s. \; ,$$

which follows by a straightforward calculation from (3.1) and the estimate

$$z_n \le 2n + \alpha - 2 + \sqrt{1 + 4(n-1)(n+\alpha-1)\cos^2\frac{\pi}{n+1}}$$

for the largest root z_n of the Laguerre polynomial $L_n^{(\alpha)}(z)$ [see Ismail and Li (1992)]. Our next result generalizes this estimate in two directions. On the one hand it also provides a lower bound, on the other hand it gives a similar approximation of every eigenvalue of the Wishart matrix.

Theorem 3.3. *Let $\lambda_1 \le \cdots \le \lambda_n$ denote the eigenvalues of the scaled Wishart matrix $\frac{1}{s}V_sV_s^T$. If $n, s(n) \to \infty$ such that for every n, $n \le s(n) \le n + K$, then for any $t \in [0, 1]$*

$$\Big|\lambda_{\lfloor nt \rfloor} - 4\cos^2\big(\frac{x_t}{2}\big)\Big| = \mathcal{O}(\sqrt{\log n/n}) \qquad a.s. \quad ,$$

where $\lambda_{\lfloor 0 \rfloor} = \lambda_1$ and x_t is the unique solution of the equation

$$x - \sin x = \pi(1 - t)$$

in the interval $[0, \pi]$. Moreover, $4\cos^2\Big(\frac{x_t}{2}\Big)$ is the t-quantile of the distribution function

$$F(z) = \frac{1}{2\pi}\int_0^z \sqrt{\frac{4-u}{u}}du$$

on the interval $[0,4]$.

Proof. Without loss of generality we may assume that $K=0$ (otherwise we consider $K+1$ subsequences seperately). If $t \in [0,1)$ it follows from Theorem 3.1 that

$$|\lambda_{\lfloor nt \rfloor} - x_{\lfloor nt \rfloor}| = \mathcal{O}(\sqrt{\log n/n}) \quad a.s.,$$

where $x_1 < \ldots < x_n$ are the roots of of the Laguerre polynomial $L_n^{(0)}(nx)$ and we use the notation $\lambda_{\lfloor 0 \rfloor} = \lambda_1$, $x_{\lfloor 0 \rfloor} = x_1$. On the other hand Theorem 8 in Gatteschi (2002) and the well known estimate

$$\Big|\frac{j_{\lfloor nt \rfloor}}{nt} - \pi\Big| = O\Big(\frac{1}{n}\Big) , \tag{3.2}$$

[see Szegö (1975), p. 15] show that the root $x_{\lfloor nt \rfloor}$ of the Laguerre polynomial $L_n^{(0)}(nx)$ can be approximated as

$$x_{\lfloor nt \rfloor} = \frac{4n+2}{n} u_{n,\lfloor nt \rfloor} + O\big(\frac{1}{n}\big) , \tag{3.3}$$

where $u_{n,\lfloor nt \rfloor} = \cos^2(U_{n,\lfloor nt \rfloor}/2)$, $U_{n,\lfloor nt \rfloor}$ is the unique solution of the equation

$$x - \sin x = \pi - \frac{4 j_{\lfloor nt \rfloor}}{4n+2} \tag{3.4}$$

in the interval $(0,\pi)$ and $j_{\lfloor nt \rfloor}$ denotes the $\lfloor nt \rfloor$th zero of the Bessel function $J_0(x)$ (the estimate is actually much sharper, but (3.3) will be sufficient for our purposes). The first assertion of the theorem now follows from (3.3), (3.2) and (3.4). The remaining statement of the Theorem is easily obtained from the representation

$$\frac{1}{2\pi} \int_0^{4\cos^2(\frac{x}{2})} \sqrt{\frac{4-u}{u}} du \;=\; \frac{\pi - x + \sin x}{\pi} ,$$

which follows by differentiating both sides with respect to the variable x. Finally, the remaining case $t=1$ is obtained by similar arguments using Theorem 9 in Gatteschi (2002). $\square$

4 Hermite ensembles

To study Hermite (or Gaussian) ensembles, or more generally, β-Hermite ensembles, we use the $n \times n$ symmetric matrix

$$G_n = \begin{bmatrix} N_1 & \frac{1}{\sqrt{2}} X_{(n-1)\beta} & & & \\ \frac{1}{\sqrt{2}} X_{(n-1)\beta} & N_2 & \frac{1}{\sqrt{2}} X_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{\sqrt{2}} X_{2\beta} & N_{n-1} & \frac{1}{\sqrt{2}} X_{\beta} \\ & & & \frac{1}{\sqrt{2}} X_{\beta} & N_n \end{bmatrix} , \tag{4.1}$$

where $\beta > 0$, and $X_\beta, \ldots, X_{(n-1)\beta}, N_1, \ldots, N_n$ are independent random variables with $X_{j\beta}^2 \sim \chi^2(j\beta)$, $X_{j\beta} \geq 0$ and $N_j \sim N(0,1)$. It was shown by Dumitriu and Edelman (2002) that the joint density of the eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ of the matrix G_n is proportional to the function defined in (1.2). Let H_n denote the nth Hermite polynomial orthogonal with respect to the weight e^{-x^2} on $\mathbb{R}$. The following result is an analogue of Theorem 2.1 for the Hermite ensemble.

Theorem 4.1. *Let $\lambda_1 \leq \cdots \leq \lambda_n$ denote the eigenvalues of the matrix (4.1) and let $x_1 < \cdots < x_n$ denote the zeros of the scaled Hermite polynomial $H_n(x/\sqrt{\beta})$. Then for every $\epsilon > 0$,*

$$P\Big\{\max_{1\leq j\leq n} |\lambda_j - x_j| \geq \epsilon\Big\} \leq 4ne^{-\epsilon^2/18}.$$

Proof. Define the non-random matrix

$$F_n = \sqrt{\frac{\beta}{2}} \begin{bmatrix} 0 & \sqrt{n-1} & & & \\ \sqrt{n-1} & 0 & \sqrt{n-2} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{2} & 0 & 1 \\ & & & 1 & 0 \end{bmatrix}.$$

It follows from the recurrence relation of the Hermite polynomials that

$$\det(xI_n - F_n) = \left(\frac{\sqrt{\beta}}{2}\right)^n H_n\left(\frac{x}{\sqrt{\beta}}\right)$$

[see Szegö (1975), page 106]. In other words: the roots $x_1, \ldots, x_n$ of the Hermite polynomial $H_n\left(x/\sqrt{\beta}\right)$ are the eigenvalues of the matrix F_n. A similar argument as given in the proof of Theorem 2.1 now shows that

$$\max_{1\leq j\leq n} |\lambda_j - x_j| \leq \|G_n - F_n\|_\infty.$$

Introducing the random variable

$$Z_n = \max\Big\{\max_{1\leq j\leq n-1} \frac{\left|X_{j\beta} - \sqrt{j\beta}\right|}{\sqrt{2}}, \max_{1\leq j\leq n} |N_j|\Big\},$$

we have $\|G_n - F_n\|_\infty \leq 3Z_n$, and it follows that

$$P\Big\{\max_{1\leq j\leq n} |\lambda_j - x_j| \geq \epsilon\Big\} \leq P\left\{Z_n \geq \frac{\epsilon}{3}\right\}. \tag{4.2}$$

By Lemma A.1 (vi) in the appendix we have

$$P\Big\{\frac{\left|X_{j\beta} - \sqrt{j\beta}\right|}{\sqrt{2}} \geq \frac{\epsilon}{3}\Big\} \leq 2e^{\psi(\sqrt{j\beta})}, \qquad j = 1, \ldots, n-1, \tag{4.3}$$

where the function ψ is defined by

$$\psi(u) = -\epsilon' u - \frac{\epsilon'^2}{2} + u^2 \log\left(1 + \frac{\epsilon'}{u}\right), \qquad \epsilon' = \frac{\sqrt{2}}{3}\epsilon.$$

To obtain an upper bound of the probability in (4.3), which does not depend on the index j, we determine $\max_{1\le j\le n-1} \psi(\sqrt{j\beta})$. For this observe first that

$$\psi'(u) = -\frac{2\epsilon' u + \epsilon'^2}{u + \epsilon'} + 2u \log\left(1 + \frac{\epsilon'}{u}\right),$$

and that for every $u > 0$,

$$\frac{(\epsilon'/u)^2}{2} = \int_0^{\epsilon'/u} x\,dx > \int_0^{\epsilon'/u} \log(1+x)\,dx = \left(1 + \frac{\epsilon'}{u}\right)\log\left(1 + \frac{\epsilon'}{u}\right) - \frac{\epsilon'}{u}.$$

This yields

$$2u \log\left(1 + \frac{\epsilon'}{u}\right) < 2u \frac{\frac{1}{2}(\epsilon'/u)^2 + (\epsilon'/u)}{1 + (\epsilon'/u)} = \frac{2\epsilon' u + \epsilon'^2}{u + \epsilon'},$$

and as a consequence $\psi'(u) < 0$, so that

$$\max_{1\le j\le n-1} \psi(\sqrt{j\beta}) = \psi(\sqrt{\beta}) \le \psi(0) = -\epsilon^2/9.$$

Hence, from (4.3),

$$P\left\{\frac{|X_{j\beta} - \sqrt{j\beta}|}{\sqrt{2}} \ge \frac{\epsilon}{3}\right\} \le 2e^{-\epsilon^2/9}, \qquad j = 1, \ldots, n-1.$$

The inequality $P\{N_1 \ge c\} \le \exp(-c^2/2)$ for $c > 0$ and Bernoulli's inequality give

$$P\left\{\max_{1\le j\le n} |N_j| \ge \frac{\epsilon}{3}\right\} = 1 - \left[P\left\{-\frac{\epsilon}{3} < N_1 < \frac{\epsilon}{3}\right\}\right]^n \le 1 - \left[1 - 2e^{-\epsilon^2/18}\right]^n \le 2ne^{-\epsilon^2/18}.$$

It now follows that

$$P\left\{Z_n \ge \frac{\epsilon}{3}\right\} \le 2(n-1)e^{-\epsilon^2/9} + 2ne^{-\epsilon^2/18} \le 4ne^{-\epsilon^2/18},$$

and an application of (4.2) yields the assertion. □

To investigate the convergence of the eigenvalues of large dimensional β-Hermite ensembles let (β_n) be a sequence of positive parameters. For every $n \ge 2$, let $\overline{\lambda}_1^{(n)} \le \cdots \le \overline{\lambda}_n^{(n)}$ be the eigenvalues of the corresponding scaled Hermite matrix $[(2n+1)\beta_n]^{-\frac{1}{2}} G_n$, and let $\xi_1^{(n)} < \cdots < \xi_n^{(n)}$ be the zeros of the scaled Hermite polynomial $H_n(\xi\sqrt{2n+1})$. Note that $\lim_{n\to\infty} \xi_n^{(n)} = 1$, see Szegö (1975), page 132.

Theorem 4.2. *Let $\overline{\lambda}_1^{(n)} \le \cdots \le \overline{\lambda}_n^{(n)}$ denote the eigenvalues of the matrix $[(2n+1)\beta_n]^{-\frac{1}{2}} G_n$ and let $\xi_1^{(n)} < \cdots < \xi_n^{(n)}$ denote the zeros of the polynomial $H_n(\xi\sqrt{2n+1})$. Then there exists an a.s. finite random variable S such that*

$$\max_{1\le j\le n} |\overline{\lambda}_j^{(n)} - \xi_j^{(n)}| \le \left\{\frac{\log n}{(2n+1)\beta_n}\right\}^{1/2} S \qquad \text{for all } n \ge 2.$$

Proof. Let (ϕ_n) be a sequence of positive numbers with $\phi_n \to \infty$ and let $\epsilon > 0$. By Theorem 4.1,

$$\begin{aligned} P\Big\{\frac{1}{\phi_n}\Big[\frac{(2n+1)\beta_n}{\log n}\Big]^{1/2} \max_{1\le j\le n}|\overline{\lambda}_j^{(n)} - \xi_j^{(n)}| \ge \epsilon\Big\} &= P\Big\{\max_{1\le j\le n}|\lambda_j^{(n)} - x_j^{(n)}| \ge \epsilon\phi_n\sqrt{\log n}\Big\} \\ &\le 4n^{1-\epsilon^2\phi_n^2/18} = o_\epsilon(n^{-2}), \end{aligned}$$

and an argument similar to that in the second part of the proof of Theorem 2.2 completes the proof. □

Remark 4.3. Dumitriu and Edelman (2004) showed the following limit assertion for the eigenvalues $\tilde\lambda_1(\beta) \le \cdots \le \tilde\lambda_n(\beta)$ of the scaled Hermite matrix $\beta^{-1/2}G_n$. For fixed dimension n and $i = 1,\ldots,n$, as $\beta \to \infty$,

$$\sqrt{\beta}\left[\tilde\lambda_i(\beta) - h_i\right] \to N \text{ in distribution,}$$

where $N \sim \mathcal{N}(0,1)$ and $h_1,\ldots,h_n$ are the zeros of the Hermite polynomial $H_n(x)$. Thus for every net (ϕ_β) with $\phi_\beta \to \infty$,

$$\frac{\sqrt{\beta}}{\phi_\beta}\left[\tilde\lambda_i(\beta) - h_i\right] \to 0 \text{ in probability.}$$

Theorem 4.1 also yields this (optimal) rate of convergence for the maximum of the absolute differences: Indeed, we have by Theorem 4.1 for every $\epsilon > 0$,

$$P\Big\{\frac{\sqrt{\beta}}{\phi_\beta}\max_{i=1}^{n}|\tilde\lambda_i(\beta) - h_i| > \epsilon\Big\} \le 4n\exp\Big(-\frac{\epsilon^2\phi_\beta^2}{18}\Big)$$

and the right hand side of this inequality converges to 0 as $\beta \to \infty$. We finally note that a similar comment can be made for the Laguerre ensemble but is omitted for the sake of brevity [see also Remark 2.3].

We conclude this section giving an analogue of Theorem 3.3.

Theorem 4.4. *Let $\overline{\lambda}_1^{(n)} \le \cdots \le \overline{\lambda}_n^{(n)}$ denote the eigenvalues of the matrix*

$$[(2n+1)\beta_n]^{-\frac{1}{2}}G_n$$

and (β_n) be a sequence of positive parameters. If $n \to \infty$, then for any $t \in [0,1]$

$$|\bar\lambda_{\lfloor nt\rfloor}^{(n)} - x_t| = \mathcal{O}\Big(\sqrt{\frac{\log n}{n\beta_n}}\Big) \quad a.s.$$

where $\bar{\lambda}^{(n)}_{\lfloor 0\rfloor} = \lambda_1$ *and* $x_t \in [-\sqrt{2}, \sqrt{2}]$ *denotes the* t*-quantile of Wigner's semi circle law, that is*

$$\frac{1}{\pi}\int_{-\sqrt{2}}^{x_t} \sqrt{2-x^2}dx \;=\; t.$$

Proof. The proof follows from Theorem 4.2 by similar arguments as presented in the proof of Theorem 3.3 observing the representations

$$\begin{aligned} H_{2m}(x) &= (-1)^m 2^{2m} m! L_m^{(-1/2)}(x^2) \\ H_{2m+1}(x) &= (-1)^m 2^{2m+1} m! x L_m^{(1/2)}(x^2) \end{aligned}$$

for the Hermite polynomials in terms of Laguerre polynomials. □

Appendix

Lemma A.1 collects some basic inequalities for random variables with a $\chi^2(r)$ and $\chi(r)$ distribution, where $r > 0$ denotes a real (not necessarily integer) parameter. These inequalities form the basis for the more intricate inequalities for the distribution of the product of the random variables $\chi(q)$ and $\chi(r)$, which will be established in Lemma A.2 below.

Lemma A.1. *Let* $r > 0$ *and* $X_r^2 \sim \chi^2(r)$, $X_r \geq 0$. *For every* $\delta > 0$,

$$\begin{aligned} &(i) && P\left\{X_r^2 \geq r+\delta\right\} \leq \Big(1+\frac{\delta}{r}\Big)^{r/2} e^{-\delta/2}, \\ &(ii) && P\left\{X_r \geq \sqrt{r}+\delta\right\} \leq \Big(1+\frac{\delta}{\sqrt{r}}\Big)^{r} \exp\Big(-\delta\sqrt{r}-\frac{\delta^2}{2}\Big), \\ &(iii) && P\left\{X_r^2 \leq r-\delta\right\} \leq \Big(1-\frac{\delta}{r}\Big)^{r/2} e^{\delta/2}, \ \textit{provided } \delta \leq r, \\ &(iv) && P\left\{X_r \leq \sqrt{r}-\delta\right\} \leq \Big(1-\frac{\delta}{\sqrt{r}}\Big)^{r} \exp\Big(\delta\sqrt{r}-\frac{\delta^2}{2}\Big), \ \textit{provided } \delta \leq \sqrt{r}. \end{aligned}$$

Moreover, for every $\delta > 0$,

$$\begin{aligned} &(v) && P\left\{\left|X_r^2 - r\right| \geq \delta\right\} \leq 2\Big(1+\frac{\delta}{r}\Big)^{r/2} e^{-\delta/2}, \\ &(vi) && P\left\{\left|X_r - \sqrt{r}\right| \geq \delta\right\} \leq 2\Big(1+\frac{\delta}{\sqrt{r}}\Big)^{r} \exp\Big(-\delta\sqrt{r}-\frac{\delta^2}{2}\Big). \end{aligned}$$

Proof. By Cramér's Theorem it follows that

$$P\left\{X_r^2 \geq r+\delta\right\} \leq \exp\left\{-\Lambda^*(r+\delta)\right\},$$

where the function Λ^* is given by

$$\Lambda^*(x) = \sup_{t<1/2} [xt - \Lambda(t)],$$

with

$$\Lambda(t) = \log E e^{tX_r^2} = -\frac{r}{2}\log(1-2t), \ t < \frac{1}{2}.$$

For $x > 0$, the supremum in the definition of Λ^* is attained at $t = (x-r)/(2x)$, which yields

$$\Lambda^*(x) = \frac{x-r}{2} + \frac{r}{2}\log\frac{r}{x},$$

$$\exp\{-\Lambda^*(r+\delta)\} = \Big(1+\frac{\delta}{r}\Big)^{r/2} e^{-\delta/2},$$

thus proving assertion (i). The proof of (iii) is similar and therefore omitted. Inequalities (ii) and (iv) are immediate consequences of (i) and (iii), respectively.
If $\delta \geq r$, then inequality (v) follows directly from (i). If $\delta < r$, then

$$\log\frac{1+(\delta/r)}{1-(\delta/r)} = 2\sum_{k=1}^{\infty}\frac{1}{2k-1}\Big(\frac{\delta}{r}\Big)^{2k-1} > 2\frac{\delta}{r},$$

and we obtain

$$\Big(1+\frac{\delta}{r}\Big)^{r/2} e^{-\delta/2} > \Big(1-\frac{\delta}{r}\Big)^{r/2} e^{\delta/2}.$$

Thus if $\delta < r$, (v) follows from (i) and (iii). Finally, inequality (vi) is derived similarly from (ii) and (iv). □

Lemma A.2. *Let $q, r > 0$ and $X_q^2 \sim \chi^2(q)$, $Y_r^2 \sim \chi^2(r)$, $X_q, Y_r \geq 0$.*

a) For every $\delta > 0$,

$$P\{X_q Y_r \geq \sqrt{qr} + \delta\} \leq \Big(1+\frac{\delta}{\sqrt{qr}}\Big)^{q/2}\exp\Big(-\frac{\delta}{2}\sqrt{\frac{q}{r}}\Big) + \Big(1+\frac{\delta}{\sqrt{qr}}\Big)^{r/2}\exp\Big(-\frac{\delta}{2}\sqrt{\frac{r}{q}}\Big).$$

b) If the random variables X_q and Y_r are independent, $q \geq r+1$, and $0 < \delta \leq \sqrt{qr}$, then

$$\begin{aligned} &P\{X_q Y_r \leq \sqrt{qr} - \delta\} \\ &\qquad \leq \Big(1-\frac{\delta}{\sqrt{qr}}\Big)^{r}\Big[\sqrt{\frac{r}{2q}}\exp\Big\{\frac{1}{2}\left(\sqrt{q}-\sqrt{r}\right)^2 + \delta\Big\} + \exp\Big(\delta\sqrt{\frac{r}{q}} - \frac{\delta^2}{2q}\Big)\Big]. \end{aligned}$$

Proof. a) We have

$$\begin{aligned} P\{X_q Y_r \geq \sqrt{qr} + \delta\} &= P\Big\{X_q^2 Y_r^2 \geq \Big(q+\delta\sqrt{\frac{q}{r}}\Big)\Big(r+\delta\sqrt{\frac{r}{q}}\Big)\Big\} \\ &\leq P\Big\{X_q^2 \geq q+\delta\sqrt{\frac{q}{r}}\Big\} + P\Big\{Y_r^2 \geq r+\delta\sqrt{\frac{r}{q}}\Big\}, \end{aligned}$$

and Lemma A.1 (i) yields the assertion.

b) Let $0 < \delta < \sqrt{qr}$. If X_q and Y_r are independent, then

$$\begin{aligned} P\{X_q Y_r \le \sqrt{qr} - \delta\} &= \int P\{X_q y \le \sqrt{qr} - \delta\}\, dP^{Y_r}(y) \\ &\le \int_{\left(\sqrt{r}-(\delta/\sqrt{q}),\infty\right)} P\{X_q y \le \sqrt{qr} - \delta\}\, dP^{Y_r}(y) \\ &\qquad + P\Big\{Y_r \le \sqrt{r} - \frac{\delta}{\sqrt{q}}\Big\} \\ &=: p_1 + p_2, \end{aligned} \tag{4.4}$$

where the last equation defines the probabilities p_1 and p_2. By Lemma A.1 (iv) we obtain

$$p_2 \le \Big(1 - \frac{\delta}{\sqrt{qr}}\Big)^r \exp\Big(\delta\sqrt{\frac{r}{q}} - \frac{\delta^2}{2q}\Big), \tag{4.5}$$

and Lemma A.1 (iii) yields

$$\begin{aligned} p_1 &= \int_{\left(\sqrt{r}-(\delta/\sqrt{q}),\infty\right)} P\Big\{X_q^2 \le q - \Big[q - \frac{(\sqrt{qr}-\delta)^2}{y^2}\Big]\Big\}\, dP^{Y_r}(y) \\ &\le \int_{\left(\sqrt{r}-(\delta/\sqrt{q}),\infty\right)} \frac{(\sqrt{qr}-\delta)^q}{q^{q/2}y^q} \exp\Big[\frac{q}{2} - \frac{(\sqrt{qr}-\delta)^2}{2y^2}\Big]\, dP^{Y_r}(y) \\ &= \frac{(\sqrt{qr}-\delta)^q e^{q/2}}{q^{q/2}2^{r/2-1}\Gamma(r/2)} \int_{\sqrt{r}-(\delta/\sqrt{q})}^{\infty} y^{r-q-1} \exp\Big[-\frac{y^2}{2} - \frac{(\sqrt{qr}-\delta)^2}{2y^2}\Big]\, dy. \end{aligned}$$

As $q \ge r+1$, y^{r-q+1} is decreasing in y, and it follows that

$$\begin{aligned} &\int_{\sqrt{r}-(\delta/\sqrt{q})}^{\infty} y^{r-q+1} y^{-2} \exp\Big[-\frac{y^2}{2} - \frac{(\sqrt{qr}-\delta)^2}{2y^2}\Big]\, dy \\ &\qquad \le \Big(\sqrt{r} - \frac{\delta}{\sqrt{q}}\Big)^{r-q+1} \int_0^{\infty} y^{-2} \exp\Big[-\frac{y^2}{2} - \frac{(\sqrt{qr}-\delta)^2}{2y^2}\Big]\, dy \\ &\qquad = \Big(\sqrt{r} - \frac{\delta}{\sqrt{q}}\Big)^{r-q+1} \frac{\sqrt{\pi}}{\sqrt{2}\,(\sqrt{qr}-\delta)} \exp(\delta - \sqrt{qr})\ , \end{aligned}$$

where the last equation can be obtained from formula (28) in Erdélyi et al. (1954), page 146. Note that Stirling's formula yields

$$\Gamma\left(\frac{r}{2}\right) \ge \sqrt{2\pi}\left(\frac{r}{2}\right)^{(r-1)/2} e^{-r/2}.$$

Combining the above inequalities, we obtain that

$$p_1 \le \sqrt{\frac{r}{2q}}\Big(1 - \frac{\delta}{\sqrt{qr}}\Big)^r \exp\Big\{\frac{1}{2}\left(\sqrt{q} - \sqrt{r}\right)^2 + \delta\Big\},$$

and the proof follows from (4.4) and (4.5). □

Lemma A.3. *Let $c_2 > 0$ and $0 \le c_1 \le c_2$. Then the function*

$$g(x) = \Big(1 - \frac{c_2}{\sqrt{(x+c_1)x}}\Big)^x,$$

defined for all $x > 0$ with $\sqrt{(x+c_1)x} \ge c_2$, is strictly increasing.

Proof. Writing

$$z = \Big(1 - \frac{c_2}{\sqrt{(x+c_1)x}}\Big)^{-1} > 1,$$

we obtain

$$\begin{aligned} \frac{d}{dx}\log g(x) &= -\log z + \Big[1 - \frac{c_1}{2(x+c_1)}\Big](z-1) \\ &\ge -\log z + \Big(1 - \frac{c_2}{2\sqrt{(x+c_1)x}}\Big)(z-1) = -\log z + \frac{1}{2}\Big(z - \frac{1}{z}\Big). \end{aligned}$$

To see that the last expression is positive note that

$$z \log z = \int_1^z 1 + \log t \, dt < \int_1^z t \, dt = \frac{1}{2}(z^2 - 1).$$

$\square$

Acknowledgements. The authors would like to thank Isolde Gottschlich, who typed parts of the paper with considerable technical expertise and Z.D. Bai and J. Silverstein for some help with the references. We are also grateful to I. Dumitriu for sending us the paper of Dumitriu and Edelman (2004) before publication. The work of the first author was supported by the Deutsche Forschungsgemeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen). Parts of this paper were written during a visit of the first author at Purdue University and this author would like to thank the Department of Statistics for its hospitality.

References

Bai, Z. D. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. Statist. Sinica 9, 611-677.

Bai, Z. D., Miao, B., and Yao, J. (2004). Convergence rates of spectral distributions of large sample covariance matrices. SIAM J. Matrix Anal. Appl., to appear.

Bai, Z. D. and Yin, Y. Q. (1988a). Convergence to the semicircle law. Ann. Probab. 16, 863-875.

Bai, Z. D. and Yin, Y. Q. (1988b). Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. Ann. Probab. 16, 1729-1741.

Bai, Z. D. and Yin, Y. Q. (1993). Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. Ann. Probab. 21, 1275-1294.

Chihara, T. S. (1978). An Introduction to Orthogonal Polynomials. Gordon and Breach, New York.

Deift, P. (1998). Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach. American Mathematical Society, Providence, Rhode Island.

Dumitriu, I. and Edelman, A. (2002). Matrix models for beta ensembles. J. Math. Phys. 43, 5830-5847.

Dumitriu, I. and Edelman, A. (2004). Eigenvalues of Hermite and Laguerre ensembles: large beta asymptotics. Preprint.

Dyson, F. J. (1962). The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics. J. Math. Phys. 3, 1199-1215.

Erdélyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F. G. (1954). Tables of Integral Transforms, Vol. I. McGraw-Hill, New York.

Fisher, R. A. (1939). The sampling distribution of some statistics obtained from nonlinear equations. Ann. Eugenics 9, 238–249.

Gatteschi, L. (2002). Asymptotics and bounds for the zeros of Laguerre polynomials: A survey. J. Comput. Appl. Math. 144, 7-27.

Horn, R. A. and Johnson, C. R. (1985). Matrix Analysis. Cambridge University Press, Cambridge.

Hsu, P. L. (1939). On the distribution of roots of certain determinantal equations. Ann. Eugenics 9, 250–258.

Ismail, M.E.H. and Li, X. (1992). Bounds on the extreme zeros of orthogonal polynomials. Proc. Amer. Math. Soc. 115, 131-140.

James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. Ann. Math. Statist. 35, 475–501.

Johnstone, I. M. (2001). On the distribution of the largest principal component. Ann. Statist. 29, 295-327.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues in certain sets of random matrices. Mat. Sb. (N.S.) 72, 507-536.

Mehta, M. L. (1967). Random Matrices. Academic Press, New York.

Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory. Wiley, New York.

Silverstein, J. W. (1985). The smallest eigenvalue of a large dimensional Wishart matrix. Ann. Probab. 13, 1364-1368.

Soshnikov, A. (2002). A note on the universality of the distribution of the largest eigenvalue in certain covariance matrices. J. Statist. Phys. 108, 1033-1056.

Szegö, G. (1975). Orthogonal Polynomials. 4th ed. Amer. Math. Soc., Providence, RI.

Tracy, C.A. and Widom, H. (2000). The distribution of the largest eigenvalue in the Gaussian ensembles: $\beta = 1, 2, 4$. Calogero-Moser-Sutherland models (Montral, QC, 1997), 461–472, CRM Ser. Math. Phys., Springer, New York, 2000.