Some asymptotic properties of the spectrum of the Jacobi ensemble

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: holger.dette@rub.de
FAX: +49 234 3214 559

Jan Nagel
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: jan.nagel@rub.de

June 20, 2008

Abstract

For the random eigenvalues with density corresponding to the Jacobi ensemble

$$c \cdot \prod_{i<j} \mid \lambda_i - \lambda_j \mid^\beta \prod_{i=1}^{n} (2-\lambda_i)^a (2+\lambda_i)^b I_{(-2,2)}(\lambda_i)$$

$(a, b > -1, \beta > 0)$ a strong uniform approximation by the roots of the Jacobi polynomials is derived if the parameters a, b, β depend on n and $n \to \infty$. Roughly speaking, the eigenvalues can be uniformly approximated by roots of Jacobi polynomials with parameters $((2a+2)/\beta - 1, (2b+2)/\beta - 1)$, where the error is of order $\{\log n/(a+b)\}^{1/4}$. These results are used to investigate the asymptotic properties of the corresponding spectral distribution if $n \to \infty$ and the parameters a, b and β vary with n. We also discuss further applications in the context of multivariate random F-matrices.

AMS Subject Classification: Primary 60F15; Secondary 15A52

Keywords and Phrases: Jacobi ensemble, strong approximations, random F-matrix, orthogonal polynomials, random matrices

1 Introduction

The three classical ensembles of random matrix theory are the Hermite, Laguerre and Jacobi ensembles. The Hermite or Gaussian ensembles arise in physics and are obtained as the eigenvalue distribution of a symmetric matrix with Gaussian entries. The Laguerre or Wishart ensembles appear in statistics and correspond to the distribution of the singular values of a Gaussian matrix. Similarly, the Jacobi ensembles are objects of statistical interest and are motivated by multivariate analysis of variance (MANOVA; see Muirhead (1982)). Associated with each ensemble there is a real positive parameter β which is usually considered for three values. The case $\beta = 1$ corresponds to real matrices, while the ensembles for $\beta = 2$ and $\beta = 4$ arise from complex and quaternion random matrices, respectively, according to Dyson's (1962) threefold classification. Dyson also observed that the eigenvalue distributions correspond to the Gibbs distribution for the classical Coloumb gas at three different temperatures. In other words – starting from the physical interpretation – for many decades it was only known that there exist random matrix models for the Coloumb gas at three different temperatures. Recently Dumitriu and Edelman (2002) provided tridiagonal random matrix models for the β-Hermite and β-Laguerre ensembles for all $\beta > 0$. Dette and Imhof (2007) derived strong uniform approximations of the random eigenvalues of these ensembles by the (deterministic) roots of Hermite and Laguerre polynomials. The development of a matrix model corresponding to the Jacobi ensemble was an open problem, which was recently considered by Lippert (2003) and Killip and Nenciu (2004) and Edelman and Sutton (2006). In particular, Killip and Nenciu (2004) found a tridiagonal matrix model for the β-Jacobi ensemble with density

$$f(\lambda) = c \prod_{i<j} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^{n} (2 - \lambda_i)^a (2 + \lambda_i)^b I_{(-2,2)}(\lambda_i), \tag{1.1}$$

$(a, b > -1, \beta > 0)$, where the entries are simple functions of independent random variables with a beta distribution on the interval $[-1, 1]$ [see equations (2.2) and (2.3) in Section 2 for more details]. It is the purpose of the present paper to obtain further insight in the stochastic properties of the random eigenvalues with density given by (1.1). In Section 2 we introduce the triangular matrix proposed by Killip and Nenciu (2004) and present a uniform approximation of the random eigenvalues with density (1.1) by roots of Jacobi polynomials. Roughly speaking, if $n \to \infty$, the random eigenvalues can be uniformly approximated by roots of the Jacobi polynomials $P_n^{((2a+2)/\beta-1,(2b+2)/\beta-1)}(x/2)$, where the error of this approximation is of order $\{\log n/(a+b)\}^{1/4}$. These results are used in Section 3 to investigate the asymptotic properties of the spectrum if the parameters a, b and β vary simultaneously with n and $n \to \infty$. In Section 4 we study as an application the eigenvalue distribution of a multivariate F-matrix which was also investigated by Silverstein (1985b) and Collins (2005). We present several extensions of these results. In particular, we consider almost sure convergence [Silverstein (1985b) discussed convergence in probability while Collins (2005) considered the expectation of the empirical spectral distribution] and investigate the case, where the parameters and the sample size converge to infinity with different order [Silverstein (1985b) and Collins (2005) discussed the case where $a \sim \gamma_1 n,\ b \sim \gamma_2 n$]. Finally, some technical results are given in an appendix.

2 Strong approximation of the Jacobi ensemble

Recall the definition of the Jacobi ensemble in (1.1) and let for $p, q > 0$ $\alpha \sim B(p,q)$ denote a Beta-distributed random variable on the interval $[-1,1]$ with density

$$2^{1-p-q}\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}(1-x)^{p-1}(1+x)^{q-1}I_{(-1,1)}(x). \tag{2.1}$$

It was shown by Killip and Nenciu (2004) that if $\alpha_0, \alpha_1, \dots, \alpha_{2n-2}$ are independent random variables with distribution

$$\alpha_k \sim \begin{cases} B(\frac{2n-k-2}{4}\beta + a + 1, \frac{2n-k-2}{4}\beta + b + 1) & \text{if } k \text{ even,} \\ B(\frac{2n-k-3}{4}\beta + a + b + 2, \frac{2n-k-1}{4}\beta) & \text{if } k \text{ odd,} \end{cases} \tag{2.2}$$

then the joint density of the (real) eigenvalues of the tridiagonal matrix

$$J := \begin{pmatrix} b_1 & a_1 & & \\ a_1 & b_2 & \ddots & \\ & \ddots & \ddots & a_{n-1} \\ & & a_{n-1} & b_n \end{pmatrix} \tag{2.3}$$

with entries

$$\begin{aligned} b_{k+1} &= (1-\alpha_{2k-1})\alpha_{2k} - (1+\alpha_{2k-1})\alpha_{2k-2}\ , \\ a_{k+1} &= \left\{(1-\alpha_{2k-1})(1-\alpha_{2k}^2)(1+\alpha_{2k+1})\right\}^{1/2} \end{aligned}$$

($\alpha_{2n-1} = \alpha_{-1} = \alpha_{-2} = -1$) is given by the Jacobi ensemble (1.1). In the following discussion we consider the asymptotic properties of the eigenvalues of the random matrix (2.3) [or equivalently of the Jacobi ensemble (1.1)], where $n \to \infty$ and the parameters a, b and β in (1.1) also vary with n. An important tool of our analysis are the Jacobi polynomials $P_n^{(\gamma,\delta)}(x)$, which are defined as the polynomials of degree n with leading coefficient $(n+\gamma+\delta+1)_n/2^n n!$ satisfying the orthogonality relation

$$\int_{-1}^{1} P_n^{(\gamma,\delta)}(x)P_m^{(\gamma,\delta)}(x)(1-x)^{\gamma}(1+x)^{\delta}dx = 0 \qquad \text{if} \quad m \neq n \tag{2.4}$$

[see Szegö (1975)]. Here and throughout this paper $(a)_n = \Gamma(a+n)/\Gamma(a)$ denotes the Pochhammer symbol. The following results provide an almost sure uniform approximation of the random eigenvalues of the Jacobi ensemble by the roots of orthogonal polynomials, if $n \to \infty$. We begin with a statement of an exponential bound for the probability of a maximal deviation between the roots of the Jacobi polynomials $P_n^{((2a+2)/\beta-1,(2b+2)/\beta-1)}(x/2)$ and random eigenvalues of the Jacobi ensemble.

Theorem 2.1. *Let $\lambda_1^{(n)} \leq \cdots \leq \lambda_n^{(n)}$ denote the ordered eigenvalues with density given by the Jacobi ensemble (1.1) and $x_1^{(n)} < \cdots < x_n^{(n)}$ denote the ordered roots of the Jacobi polynomial*

$$P_n^{(\frac{2a+2}{\beta}-1, \frac{2b+2}{\beta}-1)}(\tfrac{1}{2}x) ,$$

then the following inequality holds for any $\varepsilon \in (0,1]$

$$P\left(\max_{1\leq j\leq n}|\lambda_j^{(n)} - x_j^{(n)}| > \varepsilon\right) \leq 4(2n-1)\exp\left\{\left(\log(1+\frac{\varepsilon^2}{648+2\varepsilon^2}) - \frac{\varepsilon^2}{648+2\varepsilon^2}\right)(a+b+2)\right\} . \tag{2.5}$$

Proof of Theorem 2.1. Interchanging the rows and columns of the matrix J defined in (2.3) it follows that the matrix

$$\tilde{J} := \begin{pmatrix} b'_1 & a'_1 & & \\ a'_1 & b'_2 & \ddots & \\ & \ddots & \ddots & a'_{n-1} \\ & & a'_{n-1} & b'_n \end{pmatrix} \tag{2.6}$$

with entries

$$\begin{aligned} b'_{k+1} &:= b_{n-k} = (1-\alpha_{2n-2k-3})\alpha_{2n-2k-2} - (1+\alpha_{2n-2k-3})\alpha_{2n-2k-4} , && (2.7)\\ a'_{k+1} &:= a_{n-k-1} = \left\{(1-\alpha_{2n-2k-5})(1-\alpha_{2n-2k-4}^2)(1+\alpha_{2n-2k-3})\right\}^{1/2} && (2.8) \end{aligned}$$

has the same eigenvalues as the matrix J. This implies that the joint density of the eigenvalues of the matrix $\tilde{J}$ is also given by the β-Jacobi ensemble defined in (1.1). We now consider the deterministic matrix

$$D := \begin{pmatrix} d_1 & c_1 & & \\ c_1 & d_2 & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ & & c_{n-1} & d_n \end{pmatrix} , \tag{2.9}$$

where we essentially replace the random variables in (2.7) and (2.8) by their corresponding expectations, that is

$$\begin{aligned} d_{k+1} &:= (1-\mathrm{E}[\alpha_{2n-2k-3}])\mathrm{E}[\alpha_{2n-2k-2}] - (1+\mathrm{E}[\alpha_{2n-2k-3}])\mathrm{E}[\alpha_{2n-2k-4}] , \\ c_{k+1} &:= \left\{(1-\mathrm{E}[\alpha_{2n-2k-5}])(1-\mathrm{E}[\alpha_{2n-2k-4}]^2)(1+\mathrm{E}[\alpha_{2n-2k-3}])\right\}^{1/2} . \end{aligned}$$

A straightforward calculation [observing that the expectation and variance of a random variable with density (2.1) are given by $(q-p)/(p+q)$ and $4pq/((p+q+1)(p+q)^2)$, respectively], yields

$$(2.10) \qquad d_{k+1} = \frac{2(\tilde{b}^2-\tilde{a}^2)}{(2k+\tilde{a}+\tilde{b})(2k+\tilde{a}+\tilde{b}+2)}\,, \qquad k=0,\ldots,n-2\,;$$

$$(2.11) \qquad c_{k+1} = \frac{4}{2k+\tilde{a}+\tilde{b}+2}\left\{\frac{(k+\tilde{a}+\tilde{b}+1)(k+\tilde{a}+1)(k+\tilde{b}+1)(k+1)}{(2k+\tilde{a}+\tilde{b}+3)(2k+\tilde{a}+\tilde{b}+1)}\right\}^{1/2}$$

$(k=0,\ldots,n-3)$ where $\tilde{a}=\frac{2a+2}{\beta}$, $\tilde{b}=\frac{2b+2}{\beta}$, and (observing that $\alpha_{-1}=\alpha_{-2}=-1$)

$$(2.12) \qquad d_n = \frac{2(\tilde{b}-\tilde{a})}{2n+\tilde{a}+\tilde{b}-2}\,,$$

$$(2.13) \qquad c_{n-1} = \frac{4}{2n+\tilde{a}+\tilde{b}-2}\left\{\frac{(n+\tilde{a}-1)(n+\tilde{b}-1)(n-1)}{(2n+\tilde{a}+\tilde{b}-3)}\right\}^{1/2}.$$

For the calculation of the eigenvalues of the matrix D we consider the determinant $\det(x\ I_n-D)$, then it follows by an expansion with respect to the last row that

$$(2.14) \qquad \det(x\ I_n-D) = (x-d_n)\ G_{n-1}(x) - c_{n-1}^2\ G_{n-2}(x)\,,$$

where the polynomials $G_0(x),\ldots,G_{n-1}(x)$ are defined recursively by the three term recurrence relation

$$(2.15) \qquad G_k(x) = (x-s_k)\ G_{k-1}(x) - r_{k-1}^2\ G_{k-2}(x)$$

with coefficients

$$(2.16) \qquad s_k = \frac{2(\tilde{b}^2-\tilde{a}^2)}{(2k+\tilde{a}+\tilde{b})(2k+\tilde{a}+\tilde{b}-2)}\,,$$

$$(2.17) \qquad r_{k-1}^2 = \frac{16(k-1)(\tilde{a}+\tilde{b}+k-1)(\tilde{a}+k-1)(\tilde{b}+k-1)}{(2k+\tilde{a}+\tilde{b}-3)(2k+\tilde{a}+\tilde{b}-2)^2(2k+\tilde{a}+\tilde{b}-1)}$$

$(k=1,\ldots,n-1;\quad G_0(x):=1,\quad G_{-1}(x):=0)$. Now a straightforward calculation and a comparison with the three term recurrence relation for the monic Jacobi polynomials [see e.g. Chihara (1978), p. 220] yields

$$(2.18) \qquad G_k(x) = 2^k\hat{P}_k^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x); \qquad k=0,\ldots,n-1,$$

where $\hat{P}_k^{(\tilde{a},\tilde{b})}(x)$ denotes the kth monic Jacobi polynomial, i.e.

$$(2.19) \qquad \hat{P}_k^{(\tilde{a},\tilde{b})}(x) = \frac{2^k\ k!}{(k+\tilde{a}+\tilde{b}+1)_k}\ P_k^{(\tilde{a},\tilde{b})}(x)\,.$$

Combining equations (2.15), (2.18) and (2.19) we obtain by a tedious but straightforward calculation

$$\begin{aligned}
\det(I_n\ x - D) &= G_n(x) + (s_n - d_n)\ G_{n-1}(x) + (r_{n-1}^2 - c_{n-1}^2)\ G_{n-2}(x) \\
&= \frac{4^n n!}{(n+\tilde{a}+\tilde{b}+1)_n} \Big\{ P_n^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x) - \frac{(\tilde{b}-\tilde{a})(2n+\tilde{a}+\tilde{b}-1)}{(2n+\tilde{a}+\tilde{b}-2)(n+\tilde{a}+\tilde{b})} P_{n-1}^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x) \\
&\qquad - \frac{(2n+\tilde{a}+\tilde{b})(n+\tilde{a}-1)(n+\tilde{b}-1)}{(2n+\tilde{a}+\tilde{b}-2)(n+\tilde{a}+\tilde{b})(n+\tilde{a}+\tilde{b}-1)} P_{n-2}^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x) \Big\} \\
&= \frac{4^n n!}{(n+\tilde{a}+\tilde{b}+1)_n} \Big\{ P_n^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x) - \frac{n+\tilde{b}}{n+\tilde{a}+\tilde{b}} P_{n-1}^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x) \\
&\qquad + \frac{(2n+\tilde{a}+\tilde{b})(n+\tilde{a}-1)}{(n+\tilde{a}+\tilde{b})(n+\tilde{a}+\tilde{b}-1)} P_{n-1}^{(\tilde{a}-1,\tilde{b})}(\tfrac{1}{2}x) \Big\} ,
\end{aligned}$$

where we have used the identity

$$(2.20)\quad (n+\tilde{b}-1)P_{n-2}^{(\tilde{a},\tilde{b})}(x) = (n+\tilde{a}+\tilde{b}-1)P_{n-1}^{(\tilde{a},\tilde{b})}(x) - (2n+\tilde{a}+\tilde{b}-2)P_{n-1}^{(\tilde{a}-1,\tilde{b})}(x)$$

in the last step [see Abramovich and Stegun (1965), equation (22.7.18)]. A further application of this identity to the second polynomial yields

$$\begin{aligned}
\det(I_n x - D) &= \frac{4^n n!}{(n+\tilde{a}+\tilde{b}+1)_n} \Big\{ \frac{2n+\tilde{a}+\tilde{b}}{n+\tilde{a}+\tilde{b}} P_n^{(\tilde{a}-1,\tilde{b})}(\tfrac{1}{2}x) \\
&\qquad + \frac{(2n+\tilde{a}+\tilde{b})(n+\tilde{a}-1)}{(n+\tilde{a}+\tilde{b})(n+\tilde{a}+\tilde{b}-1)} P_{n-1}^{(\tilde{a}-1,\tilde{b})}(\tfrac{1}{2}x) \Big\} \\
&= \frac{4^n n!}{(n+\tilde{a}+\tilde{b}+1)_n} \Big\{ \frac{2n+\tilde{a}+\tilde{b}}{n+\tilde{a}+\tilde{b}} P_n^{(\tilde{a}-1,\tilde{b})}(\tfrac{1}{2}x) - \frac{2n+\tilde{a}+\tilde{b}}{n+\tilde{a}+\tilde{b}} P_n^{(\tilde{a}-1,\tilde{b})}(\tfrac{1}{2}x) \\
&\qquad + \frac{(2n+\tilde{a}+\tilde{b})(2n+\tilde{a}+\tilde{b}-1)}{(n+\tilde{a}+\tilde{b})(n+\tilde{a}+\tilde{b}-1)} P_n^{(\tilde{a}-1,\tilde{b}-1)}(\tfrac{1}{2}x) \Big\} \\
&= \frac{4^n n!}{(n+(\tilde{a}-1)+(\tilde{b}-1)+1)_n} P_n^{(\tilde{a}-1,\tilde{b}-1)}(\tfrac{1}{2}x) ,
\end{aligned}$$

where we have used the identity

$$(2.21)\quad (n+\tilde{a}-1)P_{n-1}^{(\tilde{a}-1,\tilde{b})}(x) = (2n+\tilde{a}+\tilde{b}-1)P_n^{(\tilde{a}-1,\tilde{b}-1)}(x) - (n+\tilde{a}+\tilde{b}-1)P_n^{(\tilde{a}-1,\tilde{b})}(x)$$

for the second equality [see Abramovich and Stegun (1965), equation (22.7.19)]. Consequently, the eigenvalues of the matrix D are given by the roots $x_1^{(n)} < \cdots < x_n^{(n)}$ of the Jacobi polynomial $P_n^{(\tilde{a}-1,\ \tilde{b}-1)}(\frac{1}{2}x)$. A similar argument as in Silverstein (1985a) now shows that

$$(2.22)\qquad \max_{1\le j\le n} |\lambda_j^{(n)} - x_j^{(n)}| \le \max_{1\le k\le n} \sum_{j=1}^{n} |\tilde{J}_{kj} - D_{kj}| \le 4\,\{3X_n\}^{1/2} + 6X_n ,$$

where the elements of the matrices $\tilde{J}$ and D are denoted by $\tilde{J}_{ij}$ and D_{ij}, respectively, and the random variable X_n is defined by

$$X_n := \max_{0 \le k \le 2n-2} |\alpha_k - \mathrm{E}[\alpha_k]| \, . \tag{2.23}$$

This implies for the probability in Theorem 2.1

$$\begin{aligned} P\left(\max_{1 \le j \le n} |\lambda_j^{(n)} - x_j^{(n)}| > \varepsilon \right) &\le P\left(2\{3X_n\}^{1/2} + 3X_n > \tfrac{\varepsilon}{2}\right) \le P\left(3\{3X_n\}^{1/2} > \tfrac{\varepsilon}{2}\right) \\ &= P\left(X_n > \tfrac{\varepsilon^2}{108}\right) \le \sum_{k=0}^{2n-2} P\left(|\alpha_k - \mathrm{E}[\alpha_k]| > \tfrac{\varepsilon^2}{108}\right) \end{aligned} \tag{2.24}$$

whenever $\varepsilon \in (0,1]$. Observing Lemma A.1 in the Appendix and that $\alpha_k = 1 - 2\beta_k$, where β_k is the corresponding Beta distribution on the interval $[0,1]$, it follows that

$$\begin{aligned} P\left(\max_{1 \le j \le n} \mid \lambda_j^{(n)} - x_j^{(n)} \mid > \varepsilon \right) &\le 4 \sum_{k=0}^{2n-2} \exp\left(c\ (\frac{2n-2-k}{2} \beta + a + b + 2) \right) \\ &\le 4(2n-1) \exp\left(c\ (a + b + 2)\right) \end{aligned}$$

where the constant $c = c(\varepsilon)$ is given by

$$c = \log\left(1 + \frac{\varepsilon^2}{648 + 2\varepsilon^2}\right) - \frac{\varepsilon^2}{648 + 2\varepsilon^2}. \tag{2.25}$$

This proves the assertion of the theorem. □

Note that the constant c in (2.25) is negative, and Theorem 2.1 therefore indicates that the random eigenvalues of the Jacobi ensemble can be approximated by the deterministic roots of the Jacobi polynomial $P_n^{(\tilde{a}-1,\, \tilde{b}-1)}(\frac{1}{2}x)$ with a high probability if $a + b$ is large. The following result makes this statement more precise and provides a strong uniform approximation of the eigenvalues of the Jacobi ensemble by the roots of the Jacobi polynomial $P_n^{(\tilde{a}-1,\, \tilde{b}-1)}(\frac{1}{2}x)$, if the parameters in (1.1) converge sufficiently fast to infinity. The proof follows by similar arguments as the proof of Theorem 2.2 in Dette and Imhof (2007) and is therefore omitted.

Theorem 2.2. *Let $\lambda_1^{(n)} \le \cdots \le \lambda_n^{(n)}$ denote the ordered random eigenvalues with density given by the Jacobi ensemble (1.1) with parameters $a = a_n, b = b_n, \beta = \beta_n$ and $x_1^{(n)} < \cdots < x_n^{(n)}$ denote the ordered roots of the Jacobi polynomial $P_n^{(\tilde{a}_n-1,\, \tilde{b}_n-1)}(\frac{1}{2}x)$ where*

$$\tilde{a}_n = \tfrac{2a_n+2}{\beta_n}, \qquad \tilde{b}_n = \tfrac{2b_n+2}{\beta_n}. \tag{2.26}$$

If

$$\lim_{n \to \infty} \frac{a_n + b_n}{\log n} = \infty \, , \tag{2.27}$$

then the inequality

$$\max_{1\le j\le n}\left|\lambda_j^{(n)} - x_j^{(n)}\right| \le \left(\frac{\log n}{a_n+b_n}\right)^{1/4} S$$

holds for all $n \ge 2$, where S denotes an a.s. finite random variable. In particular, if

$$\liminf_{n\to\infty}\frac{a_n+b_n}{n} > 0, \tag{2.28}$$

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

$$\max_{1\le j\le n}\left|\lambda_j^{(n)} - x_j^{(n)}\right| \le \left(\frac{\log n}{n}\right)^{1/4} S'$$

holds a.s. for all $n \ge 2$.

3 Asymptotic spectral properties

In this section we will apply Theorem 2.2 to derive the asymptotic properties of the empirical spectral distribution

$$F_n^J(\xi) := \frac{1}{n}\sum_{i=1}^n I\{\lambda_i^{(n)} \le \xi\}\ , \tag{3.1}$$

where $\lambda_1^{(n)} \le \cdots \le \lambda_n^{(n)}$ denote the ordered eigenvalues of the Jacobi ensemble defined by (1.1) with parameters $a = a_n, b = b_n$ and $\beta = \beta_n$. The results of Section 2 indicate that the empirical distribution function in (3.1) should exhibit a similar asymptotic behaviour as the empirical distribution function

$$F_n^P(\xi) := \frac{1}{n}\sum_{i=1}^n I\{x_i^{(n)} \le \xi\} \tag{3.2}$$

of the ordered roots $x_1^{(n)} < \cdots < x_n^{(n)}$ of the Jacobi polynomial $P_n^{(\tilde a_n-1,\ \tilde b_n-1)}(\frac{1}{2}x)$, where $\tilde a_n = (2a_n+2)/\beta_n,\ \tilde b_n = (2b_n+2)/\beta_n$. The asymptotic zero distribution of Jacobi polynomials has been studied by several authors [see e.g. Gawronski and Shawyer (1991), Elbert, Laforgia and Rodono (1994), Dette and Studden (1995) or Kuijlaars and Van Assche (1999) among many others], and we can use these results and Theorem 2.2 to derive the asymptotic properties of the spectrum of the Jacobi ensemble. The following result makes this statement more precise.

Theorem 3.1. *Consider the empirical distribution functions of the eigenvalues of the Jacobi ensemble (1.1) and the roots of the Jacobi polynomial $P_n^{(\tilde a_n-1,\tilde b_n-1)}(x/2)$ defined by (3.1) and (3.2), respectively, and let $(\delta_n)_{n\in\mathbb{N}},\ (\varepsilon_n)_{n\in\mathbb{N}}$ denote real sequences with $\delta_n > 0$ such that the limit*

$$\lim_{n\to\infty} F_n^P(\delta_n\xi + \varepsilon_n) = F(\xi)$$

exists at every coninuity point ξ of F. If the conditions

$$\text{(3.3)} \qquad \frac{a_n + b_n}{\log n} \xrightarrow[n\to\infty]{} \infty \ , \qquad \frac{\delta_n^4(a_n + b_n)}{\log n} \xrightarrow[n\to\infty]{} \infty$$

are satisfied, then

$$\lim_{n\to\infty} F_n^J(\delta_n \xi + \varepsilon_n) = F(\xi)$$

almost surely at every continuity point ξ of F.

Proof. Let G_n^P and G_n^J denote the empirical distribution functions of the rescaled roots

$$\frac{x_1^{(n)} - \varepsilon_n}{\delta_n}, \dots, \frac{x_n^{(n)} - \varepsilon_n}{\delta_n}$$

corresponding to the Jacobi polynomial $P_n^{(\tilde{a}_n - 1, \tilde{b}_n - 1)}(\frac{1}{2}x)$ and of the eigenvalues

$$\frac{\lambda_1^{(n)} - \varepsilon_n}{\delta_n}, \dots, \frac{\lambda_n^{(n)} - \varepsilon_n}{\delta_n}$$

corresponding to the Jacobi ensemble defined by (1.1), respectively. The arguments presented in the proof of Theorem 2.1 show that G_n^P and G_n^J are the empirical distribution functions of the eigenvalues of the matrices

$$\begin{aligned} A_n &:= \frac{1}{\delta_n}(D - \epsilon_n I_n), \\ B_n &:= \frac{1}{\delta_n}(J - \epsilon_n I_n), \end{aligned}$$

respectively. Observing Lemma 2.3 in Bai (1999) we obtain for the Levy-distance between the distribution functions G_n^P and G_n^J

$$L^3(G_n^P, G_n^J) \le \delta_n^{-2} \frac{1}{n} \sum_{i=1}^{n} \mid \lambda_i^{(n)} - x_i^{(n)} \mid^2 \ \le \delta_n^{-2} (\max_{1\le i\le n} |\lambda_i^{(n)} - x_i^{(n)}|)^2 \le S\Big(\Big(\frac{\log n}{(a_n + b_n)\delta_n^4}\Big)^{1/2}\Big) \ ,$$

where S denotes an a.s. finite random variable. Consequently, we obtain from the assumptions

$$L(G_n^P, G_n^J) \xrightarrow[n\to\infty]{a.s.} 0 \ ,$$

and the assertion of Theorem 3.1 follows observing the identities

$$\begin{aligned} G_n^P(\xi) &= F_n^P(\delta_n \xi + \varepsilon_n) \ , \\ G_n^J(\xi) &= F_n^J(\delta_n \xi + \varepsilon_n) \ . \end{aligned}$$

$\square$

Theorem 3.2. *Let $\lambda_1^{(n)} \leq \cdots \leq \lambda_n^{(n)}$ denote the ordered random eigenvalues with density given by the Jacobi ensemble (1.1). If the assumptions of Theorem 3.1 are satisfied and that there exist constants $a_1, a_2 \in \mathbb{R} \quad b_1, b_2 \in \mathbb{R}^+$ such that*

$$(i) \qquad \lim_{n\to\infty} \frac{1}{\delta_n}\left(\frac{n+\tilde{b}_n-1}{2n+\tilde{a}_n+\tilde{b}_n-2} - \varepsilon_n\right) = \frac{a_1}{2} \ ;$$

$$(ii) \qquad \lim_{n\to\infty} \frac{1}{\delta_n}\left(\frac{n(n+\tilde{a}_n-1)+(n+\tilde{b}_n-1)(n+\tilde{a}_n+\tilde{b}_n-2)}{(2n+\tilde{a}_n+\tilde{b}_n-2)^2} - \varepsilon_n\right) = \frac{a_2}{2} \ ;$$

$$(iii) \qquad \lim_{n\to\infty} \frac{1}{\delta_n^2}\frac{(n+\tilde{b}_n-1)(n+\tilde{a}_n-1)n}{(2n+\tilde{a}_n+\tilde{b}_n-2)^3} = \frac{b_1}{4} \ ;$$

$$(iv) \qquad \lim_{n\to\infty} \frac{1}{\delta_n^2}\frac{(n+\tilde{b}_n-1)(n+\tilde{a}_n-1)(n+\tilde{a}_n+\tilde{b}_n-2)n}{(2n+\tilde{a}_n+\tilde{b}_n-2)^4} = \frac{b_2}{4} \ ;$$

then the empirical distribution of the scaled eigenvalues

$$\frac{\lambda_1^{(n)} - 2(2\varepsilon_n - 1)}{2\delta_n}, \ \ldots, \ \frac{\lambda_n^{(n)} - 2(2\varepsilon_n - 1)}{2\delta_n}$$

converges almost surely to a non-degenerate distribution function, i.e.

$$\lim_{n\to\infty} F_n^J(2\delta_n\xi + 2(2\varepsilon_n - 1)) \stackrel{a.s.}{=} \int_{a_2-2\sqrt{b_2}}^{\xi} f^{(a_1,a_2,b_1,b_2)}(x)dx \ ,$$

where

$$f^{(a_1,a_2,b_1,b_2)}(x) = \begin{cases} \frac{b_1}{2\pi}\frac{\sqrt{4b_2-(x-a_2)^2}}{(b_2-b_1)x^2+(b_1a_2+b_1a_1-2b_2a_1)x+b_2a_1^2-a_1a_2b_1+b_1^2} & \text{if } \mid x-a_2 \mid \leq 2\sqrt{b_2} \ , \\ 0 & \text{else.} \end{cases}$$

Proof. By Theorem 2.2 in Dette and Studden (1995) it follows that the empirical distribution of the roots of the Jacobi polynomial $P_n^{(\tilde{a}_n-1,\ \tilde{b}_n-1)}(\frac{1}{2}x)$ has a non-degenerate limit, that is

$$F_n^P(2\delta_n\xi + 2(2\varepsilon_n - 1)) \xrightarrow[n\to\infty]{} \int_{a_2-2\sqrt{b_2}}^{\xi} f^{(a_1,a_2,b_1,b_2)}(x)dx \ .$$

The assertion is now an immediate consequence of Theorem 3.1 □

Example 3.3. Assume that

$$\lim_{n\to\infty}\frac{\tilde{a}_n}{n} = \alpha_0 \qquad \text{and} \qquad \lim_{n\to\infty}\frac{\tilde{b}_n}{n} = \beta_0$$

for some constants $\alpha_0, \beta_0 \geq 0$. If additionally $(a_n + b_n)/\log n \to \infty$, it is easy to see that the assumptions of Theorem 3.2 are satisfied with $\delta_n = \frac{1}{2}$, $\varepsilon_n = \frac{1}{2}$ $(n \in \mathbb{N})$. Consequently, it follows that the empirical spectral distribution of the Jacobi ensemble (1.1) converges almost surely to a distribution function with density

$$f_{\alpha_0,\beta_0}(x) = \frac{2+\alpha_0+\beta_0}{2\pi} \frac{\sqrt{(2r_2-x)(x-2r_1)}}{4-x^2} I_{(2r_1,2r_2)}(x) ,$$

where

$$r_1 := \frac{\beta_0^2-\alpha_0^2-4\sqrt{(\alpha_0+1)(\beta_0+1)(\alpha_0+\beta_0+1)}}{(2+\alpha_0+\beta_0)^2} ,$$
$$r_2 := \frac{\beta_0^2-\alpha_0^2+4\sqrt{(\alpha_0+1)(\beta_0+1)(\alpha_0+\beta_0+1)}}{(2+\alpha_0+\beta_0)^2} .$$

For example, if $a_n = b_n = 3n$ and $\beta_n = 2$ we have

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{n} = \lim_{n\to\infty} \frac{\tilde{b}_n}{n} = 3 , \tag{3.4}$$

and the limiting spectral distribution has the density

$$f_{3,3}(x) = \frac{4}{\pi} \frac{\sqrt{7/4-x^2}}{4-x^2} I_{[-\frac{\sqrt{7}}{2},\frac{\sqrt{7}}{2}]}(x) ,$$

which is depicted in the left part of Figure 1. Similarly, if $a_n = b_n = \sqrt{n}$ and $\beta_n = 2n$ we have

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{n} = \lim_{n\to\infty} \frac{\tilde{b}_n}{n} = 0 ,$$

and the limiting distribution is given by the arc-sine law on the interval $[-2, 2]$ with density

$$f_{0,0}(x) = \frac{1}{\pi} \frac{1}{\sqrt{4-x^2}} I_{(-2,2)}(x) ,$$

displayed in the right part of Figure 1.

Example 3.4. If the parameters $\tilde{a}_n$ and $\tilde{b}_n$ converge to infinity such that

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{n} = \infty, \qquad \lim_{n\to\infty} \frac{\tilde{b}_n}{n} = \infty, \qquad \lim_{n\to\infty} \frac{\tilde{a}_n}{\tilde{b}_n} = \gamma > 0 ,$$

and additionally

$$\beta_n \sim C \qquad \tilde{a}_n = O\left(n^{1+\nu}\right) , \tag{3.5}$$

for some $C \in \mathbb{R}^+$, $\nu \in (0,1)$ it follows that the scaled empirical distribution of the eigenvalues of the Jacobi ensemble (1.1) with parameters a_n, b_n and β_n converges almost surely to Wigner's semi circle law, that is

$$\lim_{n\to\infty} F_n^J \left(2\sqrt{\frac{n}{\tilde{a}_n-1}}\xi - 2\frac{\tilde{a}_n-\tilde{b}_n}{\tilde{a}_n+\tilde{b}_n-2} \right) \overset{a.s.}{=} \frac{2}{\pi\sigma^2} \int_{-\sigma}^{\xi} \sqrt{\sigma^2-x^2}dx \qquad |\, \xi \,| \le \sigma ,$$

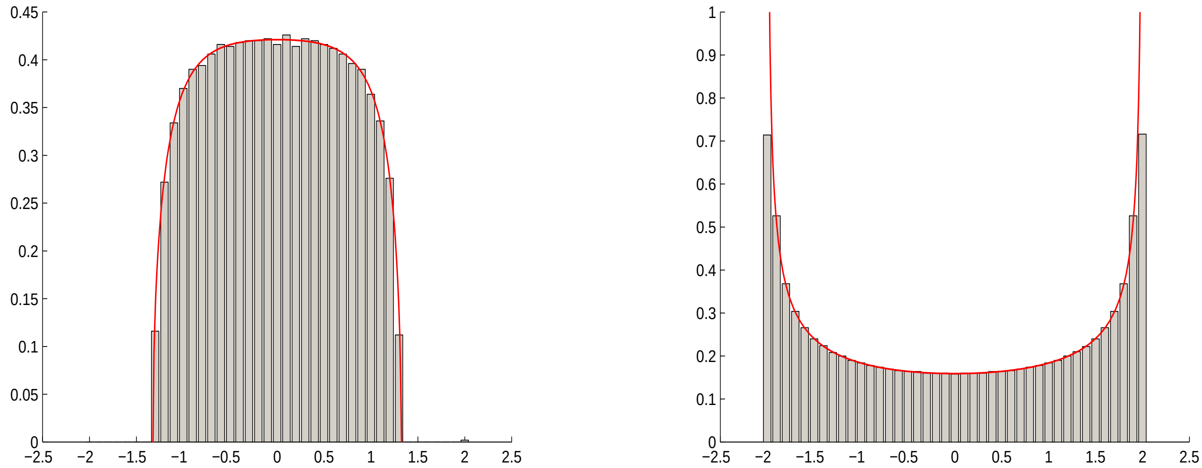

Figure 1: *Density of the limiting distribution of the random eigenvalues corresponding to the Jacobi ensemble and a histogram based on* $n = 5000$ *eigenvalues from the Jacobi ensemble (1.1). Left panel:* $a_n = b_n = 3n,\ \beta_n = 2$*; right panel:* $a_n = b_n = \sqrt{n},\ \beta_n = 2n$.

where $\sigma = 4\gamma/(1+\gamma)^{3/2}$. Note that assumption (3.5) guarantees that the second condition (3.3) in Theorem 3.1 is satisfied, i.e.

$$\lim_{n\to\infty} \sqrt[4]{\frac{n}{\tilde{a}_n - 1}}\, \frac{a_n + b_n}{\log n} = \infty.$$

This situation is also of particular interest in the case, where the inverse temperature β_n converges to 0. For example, if $a_n = b_n = n - 1$ and $\beta_n = 2n^{-1/4}$ we have $\tilde{a}_n = \tilde{b}_n = n^{5/4}$ and the condition (3.3) is satisfied. The limiting distribution of the scaled eigenvalues

$$\frac{1}{2}\sqrt{\frac{n^{5/4}-1}{n}}\lambda_1^{(n)} < \cdots < \frac{1}{2}\sqrt{\frac{n^{5/4}-1}{n}}\ \lambda_n^{(n)}$$

is then given by

$$f(x) = \frac{1}{\pi}\sqrt{2 - x^2}\ I_{[-\sqrt{2},\sqrt{2}]}(x).$$

Example 3.5. We now consider the case, where the sequences $\tilde{a}_n$ and $\tilde{b}_n$ converge to infinity with different rates. If

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{n} = \infty\ , \qquad \lim_{n\to\infty} \frac{\tilde{b}_n}{n} = \beta_0 \geq 0\,,$$

and the sequences $(\beta_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$ satisfy

$$\beta_n \sim C, \qquad \tilde{a}_n = O\left(n^{1+\nu}\right)$$

for constants $C \in \mathbb{R}^+$ and $\nu \in (0, 1/3)$, then it is easy to see that the conditions of Theorem 3.2 are satisfied with

$$\varepsilon_n = 0\,; \quad \delta_n = \frac{n}{\tilde{a}_n - 1},$$

and it follows by a straightforward calculation that

$$\lim_{n\to\infty} F_n^J\left(\frac{2n}{\tilde{a}_n - 1}\xi - 2\right) \overset{a.s.}{=} \frac{1}{4\pi}\int_{s_1}^{\xi} \frac{\sqrt{(s_2 - x)(x - s_1)}}{x}dx \qquad s_1 \le \xi \le s_2\,,$$

where

$$s_1 := 2(2 + \beta_0) - 4\sqrt{1 + \beta_0}\,,$$
$$s_2 := 2(2 + \beta_0) + 4\sqrt{1 + \beta_0}\,.$$

Similarly, if

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{n} = \infty, \qquad \lim_{n\to\infty} \frac{\tilde{b}_n}{n} = \infty, \qquad \lim_{n\to\infty} \frac{\tilde{a}_n}{\tilde{b}_n} = \infty\,,$$

and for some $\nu \in (0, 1)$ there exist constants $C, C_1, C_2 > 0$ and a $\mu > 3/2\nu - 1/2$ such that

$$\beta_n \sim C, \quad \tilde{a}_n \le C_1 n^{1+\nu}, \qquad \tilde{b}_n \ge C_2 n^{1+\mu}\,, \tag{3.6}$$

then we obtain for any $-2 \le \xi \le 6$

$$\lim_{n\to\infty} F_n^J\left(2\frac{\sqrt{n(\tilde{b}_n - 1)}}{\tilde{a}_n - 1}\xi - 2\frac{\tilde{a}_n + 2\sqrt{n(\tilde{b}_n - 1)} - \tilde{b}_n}{2n + \tilde{a}_n + \tilde{b}_n - 2}\right) \overset{a.s.}{=} \frac{1}{8\pi}\int_{-2}^{\xi} \sqrt{(6 - x)(x + 2)}dx.$$

4 An application to the multivariate F distribution

Let X_{ij} $(i = 1, \ldots, n;\ j = 1, \ldots, n_1;\ n_1 \ge n)$ and Y_{ij} $(i = 1, \ldots, n;\ j = 1, \ldots, n_2;\ n_2 \ge n)$ denote independent standard normal distributed random variables and consider the random matrices

$$X_n = (X_{ij})_{i=1,\ldots,n}^{j=1,\ldots,n_1} \in \mathbb{R}^{n\times n_1}\,; \qquad Y_n = (Y_{ij})_{i=1,\ldots,n}^{j=1,\ldots,n_2} \in \mathbb{R}^{n\times n_2}\,. \tag{4.1}$$

The matrix

$$F_n := \left(\frac{1}{n_1}X_n X_n^T\right)\left(\frac{1}{n_2}Y_n Y_n^T\right)^{-1} \in \mathbb{R}^{n\times n} \tag{4.2}$$

is called multivariate F-matrix and plays a prominent role in the multivariate analysis of variance [see e.g. Muirhead (1982)]. Silverstein (1985b) showed that under the conditions

$$\lim_{n\to\infty} \frac{n}{n_1} = y > 0 \qquad \lim_{n\to\infty} \frac{n}{n_2} = y' \in (0, 1)$$

the empirical distribution of the eigenvalues of a multivariate F-matrix converges in probability to a non-random distribution function, say $F_{y,y'}$, with density

$$f_{y,y'}(x) = \frac{1-y'}{2\pi x(xy'+y)}\sqrt{(x-s_1)(s_2-x)}I_{(s_1,s_2)}(x)\ , \tag{4.3}$$

where the bounds of the support are given by

$$\begin{aligned} s_1 &= \Big(\frac{1-\sqrt{1-(1-y)(1-y')}}{1-y'}\Big)^2\ , \\ s_2 &= \Big(\frac{1+\sqrt{1-(1-y)(1-y')}}{1-y'}\Big)^2\ . \end{aligned}$$

Moreover, if $y > 1$ the limiting distribution has mass $1 - 1/y$ at the point 0. A corresponding result for the expectation of spectral distribution of the matrix F_n can be found in Collins (2005). In the following discussion we will extend these results in two different directions using the methodology developed in Section 2. On the one hand, we prove that in the case $y \in (0, 1]$ these results are also correct, if almost sure convergence is considered. On the other hand, we extend these results to the case where n, n_1, n_2 are not necessarily of the same order.

Theorem 4.1. *Consider the multivariate F-matrix defined by (4.2). If*

$$\lim_{n\to\infty}\frac{n}{n_1} = y \in (0,1]\ , \qquad \lim_{n\to\infty}\frac{n}{n_2} = y' \in (0,1), \tag{4.4}$$

then the empirical distribution function of the eigenvalues of the matrix F_n converges a.s. to to a distribution function with density (4.3).

Proof. Consider the matrix

$$A_n := 2\left(Y_nY_n^T - X_nX_n^T\right)\left(Y_nY_n^T + X_nX_n^T\right)^{-1}\ , \tag{4.5}$$

then it follows from Muirhead (1982) that the joint density of the eigenvalues of A_n is given by the Jacobi ensemble (1.1) with $\beta = 1$ and

$$\begin{aligned} a_n &= \frac{1}{2}(n_1 - n - 1)\ , \\ b_n &= \frac{1}{2}(n_2 - n - 1)\ . \end{aligned}$$

If λ^F denotes an eigenvalue of the matrix F_n we obtain with some appropriate constant $C \in \mathbb{R}$ the identity

$$\begin{aligned} &\det\left(2\frac{\frac{n_2}{n_1}-\lambda^F}{\frac{n_2}{n_1}+\lambda^F}I_n - 2\left(Y_nY_n^T - X_nX_n^T\right)\left(Y_nY_n^T + X_nX_n^T\right)^{-1}\right) \\ &= \det\left(\lambda^F I_n - \frac{n_2}{n_1}\left(X_nX_n^T\right)\left(Y_nY_n^T\right)^{-1}\right)\cdot\det\left(Y_nY_n^T\right)\cdot C \ = \ 0\ , \end{aligned}$$

which shows that

$$\lambda^J = 2\,\frac{\frac{n_2}{n_1} - \lambda^F}{\frac{n_2}{n_1} + \lambda^F} \tag{4.6}$$

is an eigenvalue of the matrix A_n. Consequently, the empirical distribution function F_n^F of the eigenvalues of the matrix F_n satisfies the relation

$$F_n^F(\xi) \stackrel{a.s.}{=} 1 - F_n^J\left(2\frac{\frac{n_2}{n_1} - \xi}{\frac{n_2}{n_1} + \xi}\right) , \qquad \forall\, \xi \geq 0 , \tag{4.7}$$

where F_n^J denotes the empirical distribution function corresponding to the Jacobi ensemble (1.1) with parameter $\beta = 1$ (note that only the case $\xi \geq 0$ is of interest here). We now use Theorem 3.2 to derive the limiting spectral distribution. For this purpose we identify

$$\tilde{a}_n - 1 = \frac{2a_n + 2}{\beta} - 1 = n_1 - n \ , \qquad \tilde{b}_n - 1 = \frac{2a_n + 2}{\beta} - 1 = n_2 - n \ ,$$

and obtain the limits

$$\begin{aligned} \lim_{n\to\infty} \frac{\tilde{a}_n}{n} &= \lim_{n\to\infty} \frac{n_1 - n + 1}{n} = y^{-1} - 1 \geq 0 \ , \\ \lim_{n\to\infty} \frac{\tilde{b}_n}{n} &= \lim_{n\to\infty} \frac{n_2 - n + 1}{n} = y'^{-1} - 1 > 0 \ . \end{aligned}$$

Moreover, the assumption

$$\lim_{n\to\infty} \frac{a_n + b_n}{\log n} = \infty$$

is obviously satisfied. From Example 3.3 it therefore follows that the empirical distribution function F_n^J of the eigenvalues of the matrix A_n converges a.s. to a distribution function F with density

$$f(x) = \frac{y^{-1} + y'^{-1}}{2\pi} \frac{\sqrt{(2r_2 - x)(x - 2r_1)}}{4 - x^2} I_{(2r_1, 2r_2)}(x) \ ,$$

where r_1 and r_2 are given by

$$\begin{aligned} r_1 &:= \frac{(y - yy')^2 - (y' - yy')^2 - 4\sqrt{y^2 y'^2 (y + y' - yy')}}{(y + y')^2} \ , \\ r_2 &:= \frac{(y - yy')^2 - (y' - yy')^2 + 4\sqrt{y^2 y'^2 (y + y' - yy')}}{(y + y')^2} \end{aligned}$$

Observing the relation (4.7) we obtain

$$F_n^F(\xi) \xrightarrow{a.s.} 1 - F\left(2\frac{y - y'\xi}{y + y'\xi}\right) \ ,$$

and the assertion of the theorem now follows by a straightforward but tedious calculation of the density of the limiting distribution. □

While the preceding theorem essentially provides an alternative proof of the results of Silverstein (1985b), the following three theorems extend Silverstein's findings to the case where $y, y' = 0$.

Theorem 4.2. *Consider the multivariate F-matrix defined in (4.2) and denote by $\lambda_1^F \leq \cdots \leq \lambda_n^F$ the corresponding eigenvalues. If*

$$\lim_{n\to\infty} \frac{n}{n_1} = 0 \ , \qquad \lim_{n\to\infty} \frac{n}{n_2} = 0 \ , \qquad \lim_{n\to\infty} \frac{n_1}{n_2} = \gamma > 0$$

and

$$n_1 = O(n^{1+\nu})$$

with $\nu \in (0,1)$, then the empirical distribution function of the transformed eigenvalues

$$\mu_i = 2\sqrt{\frac{n_1}{n} - 1} \left\{ \frac{n_2 - n}{n_1 + n_2 - 2n} - \frac{n_2}{n_1 \lambda_i^F + n_2} \right\} \qquad i = 1, \dots, n$$

converges a.s. to a distribution function with density

$$f_\gamma(x) \ = \ \frac{2}{\pi\sigma^2} \sqrt{\sigma^2 - x^2} \ I\{-\sigma < x < \sigma\} \ ,$$

where $\sigma = 4\gamma/(1+\gamma)^{3/2}$.

Proof. Recall the definition of the matrix A_n in (4.5), which corresponds to the Jacobi ensemble (1.1) with $\beta = 1$, $a_n = \frac{1}{2}(n_1 - n - 1)$, $b_n = \frac{1}{2}(n_2 - n - 1)$. Using the notation (2.26) we obtain $\tilde{a}_n - 1 = n_1 - n$, $\tilde{b}_n - 1 = n_2 - n$. By the assumption of the theorem we have

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{n} = \infty \ , \qquad \lim_{n\to\infty} \frac{\tilde{b}_n}{n} = \infty \ , \qquad \lim_{n\to\infty} \frac{\tilde{a}_n}{\tilde{b}_n} = \gamma$$

and $\tilde{a}_n = O(n^{1+\nu})$. Therefore it follows from Example 3.4 that

$$\begin{aligned} &\lim_{n\to\infty} F_n^J \left(2\sqrt{\frac{n}{n_1 - n}}\xi - 2\frac{n_1 - n_2}{n_1 + n_2 - 2n} \right) \\ =&\frac{2}{\pi\sigma^2} \int_\infty^\xi \sqrt{\sigma^2 - x^2} I\{-\sigma < x < \sigma\} \, dx \end{aligned}$$

a.s., where F_n^J denotes the empirical distribution of the eigenvalues of the matrix A_n. The identity (4.7) implies for $\xi > -2$

$$F_n^J(\xi) \stackrel{a.s.}{=} 1 - F_n^F \left(\frac{n_2}{n_1} \frac{2-\xi}{2+\xi} \right)$$

and therefore it follows

$$\lim_{n\to\infty} F_n^F\left(\frac{n_2}{n_1}\left(\frac{2}{2\frac{n_2-n}{n_1+n_2-2n}-\sqrt{\frac{n}{n_1-n}}\xi}-1\right)\right)=\frac{2}{\pi\sigma^2}\int_{\infty}^{\xi}\sqrt{\sigma^2-x^2}I\left\{-\sigma<x<\sigma\right\}dx$$

a.s., which proves the assertion of the theorem. $\square$

Theorem 4.3. *Consider the multivariate F-matrix and denote by $\lambda_1^F \le \cdots \le \lambda_n^F$ the corresponding eigenvalues. If*

$$\lim_{n\to\infty}\frac{n}{n_1}=0\ ,\qquad \lim_{n\to\infty}\frac{n}{n_2}=y'\in(0,1]$$

and

$$n_1=O(n^{1+\nu})\ ,$$

with $\nu\in(0,1/3)$, then the empirical distribution function of the scaled eigenvalues

$$\mu_i=\frac{n}{2(n_1-n)}\left(\lambda_i^F\frac{n_1}{n_2}+1\right)\qquad i=1,\dots,n$$

converges a.s. to a distribution function F with density

$$f_{y'}(x)=\frac{1}{4\pi}\frac{\sqrt{(xs_2-1)(1-xs_1)}}{x^2}I_{(s_2^{-1},s_1^{-1})}(x)\ ,$$

where the bounds of the support of the density are given by

$$\begin{aligned}s_1&:=2(y'^{-1}+1)-4\sqrt{y'^{-1}}\ ,\\ s_2&:=2(y'^{-1}+1)+4\sqrt{y'^{-1}}\ .\end{aligned}$$

The proof is analogous to the proof of Theorem 4.2, using the first result of Example 3.5. Similarly, the following theorem can be proven using the last statement in Example 3.5.

Theorem 4.4. *Consider the multivariate F-matrix and denote by $\lambda_1^F \le \cdots \le \lambda_n^F$ the corresponding eigenvalues. If*

$$\lim_{n\to\infty}\frac{n}{n_1}=0\ ,\qquad \lim_{n\to\infty}\frac{n}{n_2}=0,\qquad \lim_{n\to\infty}\frac{n_1}{n_2}=\infty$$

and for some $\nu\in(0,1)$ there exist constants $C_1,C_2>0$ and a $\mu>3/2\nu-1/2$ such that

$$n_1\le C_1n^{1+\nu},\qquad n_2\ge C_2n^{1+\mu},$$

then the empirical distribution function of the scaled eigenvalues

$$\mu_i = 2\frac{n_1 - n}{n_1 + n_2}\frac{\frac{n_1}{n_2}(n_2 - \sqrt{n(n_2-n)})\lambda_i^F - (n_1 + \sqrt{n(n_2-n)})}{\sqrt{n(n_2-n)}(1+\frac{n_1}{n_2}\lambda_i^F)} \qquad i = 1, \ldots, n$$

converges a.s. to a distribution function F with density

$$f(x) = \frac{1}{8\pi}\sqrt{(6+x)(2-x)}I_{[-6,2]}(x) \ .$$

5 Appendix: auxiliary results

Lemma A.1. *Let Z denote a Beta-distributed random variable on the interval $[0,1]$ with density*

$$\text{(5.1)} \qquad \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\ x^{p-1}(1-x)^{q-1}\ I_{(0,1)}(x) \qquad (p, q > 0)\ ,$$

then for any $\delta > 0$

$$P(\mid Z - E[Z] \mid > \delta) \leq 4e^{c(p+q)}\ ,$$

where the constant c is defined by

$$c = \log\left(1 + \frac{\delta}{3+2\delta}\right) - \frac{\delta}{3+2\delta}\ .$$

Proof. If $X \sim \Gamma(p,\ p+q),\ Y \sim \Gamma(q,\ p+q)$ denote independent Gamma-distributed random variables, it is well known that the ratio $Z = X/(X+Y)$ has a Beta-distribution with density (5.1). Because $E[Z] = E[X] = p/(p+q)$ it follows that

$$\text{(5.2)} \qquad P(|Z - \mathrm{E}[Z]| > \delta) = P\Big(\Big|\frac{X}{X+Y} - \mathrm{E}[X]\Big| > \delta\Big)$$

Define $\delta' = \delta/(3+2\delta)$ and assume that

$$\mid X - E[X] \mid \leq \delta'\ , \quad \mid Y - E[Y] \mid \leq \delta'\ ,$$

then it is easy to see that $\mid X+Y-1 \mid = \mid X+Y-E[X+Y] \mid \leq 2\delta'$; and

$$\Big|\frac{X}{X+Y} - \mathrm{E}[X]\Big| \leq \frac{1}{1-2\delta'}(|X - \mathrm{E}[X]| + \mathrm{E}[X]|X+Y-\mathrm{E}[X+Y]|) \leq \frac{1}{1-2\delta'}3\delta' = \delta\ .$$

This implies for the probability in (5.2)

$$\text{(5.3)} \quad P\left(\Big|\frac{X}{X+Y} - \mathrm{E}[X]\Big| > \delta\right) \ \leq\ P(X > \mathrm{E}[X]+\delta') + P(X < \mathrm{E}[X]-\delta') + P(Y > \mathrm{E}[Y]+\delta') + P(Y < \mathrm{E}[Y]+\delta')\ .$$

Using similar arguments as in Dette and Imhof (2007) we obtain the estimates

$$\begin{aligned} P(U > E[U] + \delta') &\leq \exp\{(p+q)(\log(1+\delta') - \delta')\} , \\ P(U < E[U] - \delta') &\leq \exp\{(p+q)(\log(1-\delta') + \delta')\} , \end{aligned}$$

where U is either X or Y. The assertion of Lemma A.1 now follows from (5.3) and the definition of δ' observing that $\log(1+\delta') - \delta' > \log(1-\delta') + \delta'$. $\square$

Acknowledgements. The authors are grateful to Martina Stein who typed parts of this paper with considerable technical expertise. The work of the authors was supported by the Sonderforschungsbereich Tr/12 (project C2, Fluctuations and universality of invariant random matrix ensembles) and in part by a NIH grant award IR01GM072876:01A1.

References

M. Abramovich, I. Stegun (1965). Handbook of Mathematical Functions. Dover Publications Inc., N.Y.

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

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

B. Collins (2005). Product of random projections, Jacobi ensembles and universality problems arising from free probability. Probability Theory and Related Fields 133, 315-344

H. Dette and L. A. Imhof (2007). Uniform approximation of eigenvalues in Laguerre and Hermite β-ensembles by roots of orthogonal polynomials. Transactions of the American Mathematical Society 359, 4999-5018.

H. Dette and W.J. Studden (1995). Some new asymptotic properties for the zeros of the Jacobi, Laguerre and Hermite polynomials. Constructive Approximation, Vol. 11, 227-238.

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

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

Edelman and Sutton (2006). The beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems. Found. Comput. Math. In press. http://dx.doi.org/10.1007/s10208-006-0215-9.

A. Elbert, A. Laforgia, L. G. Rodono (1994). On the zeros of Jacobi polynomials. Acta Math. Hungar. 64, 351-359.

W. Gawronski, B. Shawyer (1991). Strong asymptotics and the limit distribution of the zeros of Jacobi polynomials $P_n^{(a_n+\alpha,b_n+\beta)}(x)$. In: Progress in Approximation Theory (P. Nevai, A. Pinkus eds.). New York: Academic Press 379-404.

R. Killip and I. Nenciu I (2004). Matrix models for circular ensembles. Int. Math. Res. Not. 50, 2665-2701.

A. B. J. Kuijlaars and W. van Assche (1999). The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients, J. Approx. Theory 99 (1999), 167-197.

R. A. Lippert (2003). A matrix model for the beta-Jacobi ensemble. J. Math. Phys. 44 4807-4816.

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

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

J. Silverstein (1985b). The limiting eigenvalue distribution of a multivariate F-matrix. SIAM J. Math. Anal. 16, 641-646.

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