Bernoulli **19**(3), 2013, 1028–1046
DOI: 10.3150/12-BEJ495

Limit theorems for beta-Jacobi ensembles

TIEFENG JIANG

School of Statistics, University of Minnesota, 224 Church Street SE, Minneapolis, MN 55455, USA.
E-mail: tjiang@stat.umn.edu

For a β-Jacobi ensemble determined by parameters a_1, a_2 and n, under the restriction that the three parameters go to infinity with n and a_1 being of small orders of a_2, we obtain some limit theorems about the eigenvalues. In particular, we derive the asymptotic distributions for the largest and the smallest eigenvalues, the central limit theorems of the eigenvalues, and the limiting distributions of the empirical distributions of the eigenvalues.

Keywords: beta-ensemble; empirical distribution; Jacobi ensemble; Laguerre ensemble; largest eigenvalue; limiting distribution; random matrix; random operator; smallest eigenvalue

1. Introduction

Let $\beta > 0$ be a constant and $n \geq 2$ be an integer. A beta-Jacobi ensemble, also called in the literature as the beta-MANOVA ensemble, is a set of random variables $(\lambda_1, \ldots, \lambda_n) \in [0, 1]^n$ with probability density function

$$f_{\beta,a_1,a_2}(\lambda) = c_J^{\beta,a_1,a_2} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta \cdot \prod_{i=1}^n \lambda_i^{a_1-p}(1-\lambda_i)^{a_2-p}, \tag{1.1}$$

where $a_1, a_2 > \frac{\beta}{2}(n-1)$ are parameters, $p = 1 + \frac{\beta}{2}(n-1)$, and

$$c_J^{\beta,a_1,a_2} = \prod_{j=1}^n \frac{\Gamma(1+\beta/2)\Gamma(a_1+a_2-(\beta/2)(n-j))}{\Gamma(1+(\beta/2)j)\Gamma(a_1-(\beta/2)(n-j))\Gamma(a_2-(\beta/2)(n-j))}. \tag{1.2}$$

The ensemble has close connections to the multivariate analysis of variance (MANOVA). For $\beta = 1, 2$ and 4, the function $f_\beta(\lambda)$ in (1.1) is the density function of the eigenvalues of $\mathbf{Y}^*\mathbf{Y}(\mathbf{Y}^*\mathbf{Y} + \mathbf{Z}^*\mathbf{Z})^{-1}$ with $a_1 = \frac{\beta}{2}m_1$ and $a_2 = \frac{\beta}{2}m_2$, where $\mathbf{Y} = \mathbf{Y}_{m_1 \times n}$ and $\mathbf{Z} = \mathbf{Z}_{m_2 \times n}$ are independent matrices with $m_1, m_2 \geq n$, and the entries of both matrices are independent random variables with the standard real, complex or quaternion Gaussian distributions. See [9] and [37] for $\beta = 1, 2$. Other references about the connections between the Jacobi ensembles and statistics are [3,7–9,12,20,24,29,30,37].

In statistical mechanics, the model of the log gases can be characterized by the beta-Jacobi ensembles. A log gas is a system of charged particles on the real line which are subject to a logarithmic interaction potential and Brownian-like fluctuations. If the particles are contained in the interval [0, 1] and are also subject to the external potential $\sum_{i=1}^n (\frac{r+1}{2} - \frac{1}{\beta}) \log \lambda_i + (\frac{s+1}{2} - \frac{1}{\beta}) \log(1-\lambda_i)$, where $r = \frac{2}{\beta}a_1 - n$ and $s = \frac{2}{\beta}a_2 - n$, and β is the inverse of the temperature, then it is known that the stationary distribution of the system of charges in the long term is the Jacobi ensemble as in (1.1), see, for example, [5,19,25,43].

The beta-Jacobi ensembles also have connections to other subjects in mathematics and physics. See, for instance, the lattice gas theory [24,25], Selberg integrals [26,36,38] and Jack functions [4,32,42].

Now we briefly recall some research on the beta-Jacobi ensembles. Lippert [35] gives a model to generate the beta-Jacobi ensembles (see also [34] for a similar method used in the construction of the beta-circular ensembles). In studying the largest principal angles between random subspaces, Absil, Edelman and Koev [1] obtain a formula related to the Jacobi ensembles. Edelman and Sutton [22] study CS decomposition and singular values about these models. Dumitriu and Koev [18] derive the exact distributions of the largest eigenvalues for the ensembles. Jiang [30] derives the bulk and the edge scaling limits for the beta-Jacobi ensembles for $\beta = 1$ and 2 when p and a_1 in (1.1) are of small orders of a_2. Johnstone [31] obtains the asymptotic distribution of the largest eigenvalues for $\beta = 1$ and 2 when a_1, a_2 and p in (1.1) are proportional to each other. Recently, Demni [10] investigates the beta-Jacobi processes.

In this paper, for the beta-Jacobi ensembles, we study the asymptotic distributions of the largest and smallest eigenvalues, the limiting empirical distributions of the eigenvalues, the law of large numbers and the central limit theorems for the eigenvalues. Before stating the main results, we need some notation.

Let $\beta > 0$ be a fixed constant, $n \ge 2$ be an integer, a_1 and a_2 be positive variables. The following condition will be used later.

$$\begin{aligned} &n \to \infty, a_1 \to \infty \text{ and } a_2 \to \infty \text{ such that} \\ &a_1 = \mathrm{o}(\sqrt{a_2}), \qquad n = \mathrm{o}(\sqrt{a_2}) \quad \text{and} \quad \frac{n\beta}{2a_1} \to \gamma \in (0, 1]. \end{aligned} \tag{1.3}$$

For two Borel probability measures μ and ν on $\mathbb{R}^k$, recall the metric

$$d(\mu, \nu) = \sup_{\|f\|_{\mathrm{BL}} \le 1} \left| \int_{\mathbb{R}^k} f(x)\,\mathrm{d}\mu - \int_{\mathbb{R}^k} f(x)\,\mathrm{d}\nu \right|, \tag{1.4}$$

where $f(x)$ is a bounded Lipschitz function defined on $\mathbb{R}^k$ with

$$\|f\|_{\mathrm{BL}} = \sup_{x \ne y} \frac{|f(x) - f(y)|}{\|x - y\|} + \sup_{x \in \mathbb{R}^k} |f(x)|.$$

Then, for a sequence of probability measures $\{\mu_n; n = 0, 1, 2, \ldots\}$ defined on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, we know μ_n converges weakly to μ_0 if and only if $d(\mu_n, \mu_0) \to 0$ as $n \to \infty$, see, for example, [14]. Similarly, we say that a sequence of random variables $\{Z_n; n \ge 1\}$ taking values in $\mathbb{R}^k$ converges weakly (or in distribution) to a Borel probability measure μ on $\mathbb{R}^k$ if $Ef(Z_n) \to \int_{\mathbb{R}^k} f(x)\mu(\mathrm{d}x)$ for any bounded and continuous function $f(x)$ defined on $\mathbb{R}^k$. This is also equivalent to that $d(\mathcal{L}(Z_n), \mu) \to 0$ as $n \to \infty$, where $\mathcal{L}(Z_n)$ is the probability distribution of Z_n, see also [14].

For $\gamma \in (0,1]$, let $\gamma_{\min} = (\sqrt{\gamma}-1)^2$ and $\gamma_{\max} = (\sqrt{\gamma}+1)^2$. The Marchenko–Pastur law is the probability distribution with density function

$$f_\gamma(x) = \begin{cases} \frac{1}{2\pi\gamma x}\sqrt{(x-\gamma_{\min})(\gamma_{\max}-x)}, & \text{if } x \in [\gamma_{\min}, \gamma_{\max}], \\ 0, & \text{otherwise.} \end{cases} \tag{1.5}$$

Our first result is the global behavior of the eigenvalues of the beta-Jacobi ensembles.

Theorem 1 (Empirical law). *Let $\lambda_1, \ldots, \lambda_n$ be random variables with density function $f_{\beta,a_1,a_2}(\lambda)$ as in (1.1). Set*

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{(a_2/n)\lambda_i}$$

for $n \ge 2$. Assuming (1.3), then $d(\mu_n, \mu_0)$ converges to zero in probability, where μ_0 has density $c \cdot f_\gamma(cx)$ with $c = 2\gamma/\beta$ and $f_\gamma(x)$ is as in (1.5).

The next result gives the scales of the largest and smallest eigenvalues for the beta-Jacobi ensembles.

Theorem 2 (Law of large numbers for extreme eigenvalues). *Let $\lambda_1, \ldots, \lambda_n$ be random variables with density function $f_{\beta,a_1,a_2}(\lambda)$ as in (1.1). Set $\lambda_{\max}(n) = \max\{\lambda_1, \ldots, \lambda_n\}$, and $\lambda_{\min}(n) = \min\{\lambda_1, \ldots, \lambda_n\}$. Assuming (1.3), we have that*

$$\frac{a_2}{n} \cdot \lambda_{\max}(n) \to \beta \cdot \frac{(1+\sqrt{\gamma})^2}{2\gamma} \quad \text{and} \quad \frac{a_2}{n} \cdot \lambda_{\min}(n) \to \beta \cdot \frac{(1-\sqrt{\gamma})^2}{2\gamma}$$

in probability.

The following is the asymptotic behavior of the "trace" of the beta-Jacobi ensembles.

Theorem 3 (Central limit theorems). *Let $\lambda_1, \ldots, \lambda_n$ be random variables with density function $f_{\beta,a_1,a_2}(\lambda)$ as in (1.1). Given integer $k \ge 1$, define*

$$X_i = \sum_{j=1}^{n}\left(\frac{ca_2}{n}\lambda_j\right)^i - n\sum_{r=0}^{i-1}\frac{1}{r+1}\binom{i}{r}\binom{i-1}{r}\gamma^r$$

for $i \ge 1$, where $c = 2\gamma/\beta$ and γ is as in (1.3). Assuming (1.3), then $(X_1, \ldots, X_k)$ converges weakly to a multivariate normal distribution $N_k(\mu, \Sigma)$ for some μ and Σ given in Theorem 1.5 *from* [17].

Killip [33] obtains the central limit theorem for $\sum_{i=1}^{n} I(\lambda_i \in (a,b))$, where $a < b$ are two constants. Theorem 3 is the central limit theorem for homogenous polynomials of λ_i's.

From the recent results by Ramírez and Rider [40] and Ramírez, Rider and Virág [39], we are able to investigate the asymptotic distributions of the smallest and largest eigenvalues for the beta-Jacobi ensembles next. Look at the operator

$$\mathcal{T}_{\beta,a} = -\exp\left[(a+1)x + \frac{2}{\sqrt{\beta}}b(x)\right]\frac{\mathrm{d}}{\mathrm{d}x}\left\{\exp\left[-ax - \frac{2}{\sqrt{\beta}}b(x)\right]\right\},$$

where $a > -1$ and $\beta > 0$ are constants, and $b(x)$ is a standard Brownian motion on $[0, \infty)$. With probability one, when restricted to the positive half-line with Dirichlet conditions at the origin, $\mathcal{T}_{\beta,a}$ has discrete spectrum comprised of simple eigenvalues $0 < \Lambda_0(\beta, a) < \Lambda_1(\beta, a) < \cdots \uparrow \infty$ as stated in Theorem 1 from [40].

For a sequence of pairwise different numbers $a_1, \ldots, a_n$, let $a^{(1)} > a^{(2)} > \cdots > a^{(n)}$ be their order statistics.

Theorem 4 (Limiting distribution of smallest eigenvalues). *Let $\lambda_1, \ldots, \lambda_n$ be random variables with density function $f_{\beta,a_1,a_2}(\lambda)$ as in (1.1). Let $c > 0$ be a constant, and $2\beta^{-1}a_1 - n = c$. If $n \to \infty$ and $a_2 \to \infty$ such that $n = \mathrm{o}(\sqrt{a_2})$, then $(2\beta^{-1}na_2) \cdot (\lambda^{(n)}, \ldots, \lambda^{(n-k+1)})$ converges weakly to $(\Lambda_0(\beta, c), \Lambda_1(\beta, c), \ldots, \Lambda_{k-1}(\beta, c))$.*

Now look at another random operator

$$-\mathcal{H}_\beta = \frac{\mathrm{d}^2}{\mathrm{d}x^2} - x - \frac{2}{\sqrt{\beta}}b'_x, \tag{1.6}$$

where b_x is a standard Brownian motion on $[0, +\infty)$ (b'_x is not the derivative of b_x since it is not differentiable almost everywhere). We use equation (1.6) in the following sense. For $\lambda \in \mathbb{R}$ and function $\psi(x)$ defined on $[0, +\infty)$ with $\psi(0) = 0$ and $\int_0^\infty ((\psi')^2 + (1+x)\psi^2)\,\mathrm{d}x < \infty$, we say (ψ, λ) is an eigenfunction/eigenvalue pair for $-\mathcal{H}_\beta$ if $\int_0^\infty \psi^2(x)\,\mathrm{d}x = 1$ and

$$\psi''(x) = \frac{2}{\sqrt{\beta}}\psi(x)b'_x + (x+\lambda)\psi(x)$$

holds in the sense of integration-by-parts, that is,

$$\psi'(x) - \psi'(0) = \frac{2}{\sqrt{\beta}}\psi(x)b_x + \int_0^x -\frac{2}{\sqrt{\beta}}b_y\psi'(y)\,\mathrm{d}y + \int_0^x (y+\lambda)\psi(y)\,\mathrm{d}y.$$

Theorem 1.1 from [40] says that, with probability one, for each $k \ge 1$, the set of eigenvalues of $-\mathcal{H}_\beta$ has well-defined k-largest eigenvalues Λ_k. Recall (1.1), set

$$m_n = \left(\sqrt{n} + \sqrt{2\beta^{-1}a_1}\right)^2 \quad \text{and} \quad \sigma_n = \frac{(2\beta^{-1}na_1)^{1/6}}{(\sqrt{n} + \sqrt{2\beta^{-1}a_1})^{4/3}}.$$

Theorem 5 (Limiting distribution of largest eigenvalues). *For each $k \ge 1$, let Λ_k be the kth largest eigenvalue of $-\mathcal{H}_\beta$ as in (1.6). Let $\lambda_1, \ldots, \lambda_n$ be random variables with joint density function $f_{\beta,a_1,a_2}(\lambda)$ as in (1.1). Assuming (1.3), then $\sigma_n((2a_2\beta^{-1})\lambda^{(l)} - m_n)_{l=1,\ldots,k}$ converges weakly to $(\Lambda_1, \ldots, \Lambda_k)$.*

Remark 1.1. Dumitriu and Koev [18] derive the exact formulas for the cumulative distribution functions of the largest and smallest eigenvalue of the beta-Jacobi ensembles as in (1.1) for fixed parameter β, a_1, a_2 and n. Here we complement their work by proving the asymptotic distributions in Theorems 4 and 5.

Remark 1.2. In [30], Jiang studies Theorems 1, 2, 3 and 5 for $\beta = 1$ and 2, which are special cases of the current theorems. The method used in [30] is the approximation of the entries of Haar-invariant orthogonal or unitary matrices by independent and identically distributed real or complex Gaussian random variables.

Remark 1.3. In the above theorems, the assumption that $a_1 = \mathrm{o}(\sqrt{a_2})$ is important. It is possible that some of the above theorems may still hold if $a_1 = \mathrm{o}(a_2)$, see, for example, Theorem 2 from [13] as $\beta = 2$ for a conclusion on the circular law. However, Theorems 1 and 5 will no longer be true if $a_1/a_2 \to c \in (0, \infty)$. See, for example, [11] in the context of Theorem 1 for any $\beta > 0$, or [31] in the context of Theorem 5 as $\beta = 1$.

In summary, we study the eigenvalues of the beta-Jacobi ensembles in this paper. We obtain the empirical law and the central limit theorems for the eigenvalues as well as the scales and limiting distributions of the largest and smallest eigenvalues. These results can be applied to various statistical, mathematical and physical problems mentioned at the beginning of this section if the three parameters n, a_1 and a_2 satisfy the restriction (1.3).

The proofs of the above theorems are based on an approximation result. In fact, we approximate the beta-Jacobi ensembles by the beta-Laguerre ensembles through measuring the variation distance between the eigenvalues in the two ensembles. Then the known results for the beta-Laguerre ensemble are used to get those for the beta-Jacobi ensembles.

Let μ and ν be probability measures on $(\mathbb{R}^m, \mathcal{B})$, where $m \geq 1$ and $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}^m$. The variation distance $\|\mu - \nu\|$ is defined by

$$\|\mu - \nu\| = 2 \cdot \sup_{A \in \mathcal{B}} \bigl|\mu(A) - \nu(A)\bigr| = \int_{\mathbb{R}^m} \bigl|f(x) - g(x)\bigr| \, \mathrm{d}x_1 \cdots \mathrm{d}x_m, \tag{1.7}$$

where $f(x)$ and $g(x)$ are the density functions of μ and ν with respect to the Lebesgue measure, respectively. For a random variable Z, we use $\mathcal{L}(Z)$ to denote its probability distribution. The following is the tool to obtain the results stated earlier.

Theorem 6. *Let $\mu = (\mu_1, \ldots, \mu_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_n)$ be random variables with density $f_{\beta,a_1}(\mu)$ as in (2.1) (taking $a = a_1$) and $f_{\beta,a_1,a_2}(\lambda)$ as in (1.1). Assuming (1.3), then $\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| \to 0$.*

Finally the outline of this paper is given as follows. In Section 2, some known conclusions and some results on the beta-Laguerre ensembles are reviewed and proved. They will be used in the proofs of the main theorems. In Section 3, the proof of Theorem 6 is presented. In Section 4, we prove Theorems 1–5.

2. Some auxiliary results on β-Laguerre ensembles

Let $\beta > 0$ be a constant, $n \geq 2$ be an integer, $p = 1 + \frac{\beta}{2}(n-1)$ and parameter $a > \frac{\beta}{2}(n-1)$. A β-Laguerre (Wishart) ensemble is a set of non-negative random variables $(\lambda_1, \ldots, \lambda_n) := \lambda$ with probability density function

$$f_{\beta,a}(\lambda) = c_L^{\beta,a} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta \cdot \prod_{i=1}^n \lambda_i^{a-p} \cdot \mathrm{e}^{-1/2 \sum_{i=1}^n \lambda_i}, \tag{2.1}$$

where

$$c_L^{\beta,a} = 2^{-na} \prod_{j=1}^n \frac{\Gamma(1+\beta/2)}{\Gamma(1+(\beta/2)j)\Gamma(a-(\beta/2)(n-j))}. \tag{2.2}$$

One can see [16] for the construction of a matrix to generate eigenvalues with such a distribution. If $\mathbf{X} = (x_{ij})$ is an $m \times n$ matrix with $m \geq n$, where x_{ij}'s are independent and identically distributed random variables with the standard real normal ($\beta = 1$), complex normal ($\beta = 2$) or quaternion normal ($\beta = 4$) distribution, then $f_\beta(\lambda)$ is the density function of the eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_n)$ of $\mathbf{X}^*\mathbf{X}$ with $a = \frac{\beta}{2}m$ for $\beta = 1, 2,$ or 4. See [23,28,37] for the cases $\beta = 1$ and 2, and [36] for $\beta = 4$, or (4.5) and (4.6) from [21].

The next lemma follows from Theorem 1.5 in [17].

Lemma 2.1. *Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be random variables with the density function as in (2.1). If $n \to \infty, a \to \infty$ and $n\beta/(2a) \to \gamma \leq 1$, then*

(i) $$\frac{1}{n^{i+1}} \sum_{j=1}^n \lambda_j^i \text{ converges to } \left(\frac{\beta}{\gamma}\right)^i \sum_{r=0}^{i-1} \frac{1}{r+1} \binom{i}{r} \binom{i-1}{r} \gamma^r \text{ in probability;}$$

(ii) $$\frac{1}{n^i} \sum_{j=1}^n \lambda_j^i - \left(\frac{\beta}{\gamma}\right)^i n \sum_{r=0}^{i-1} \frac{1}{r+1} \binom{i}{r} \binom{i-1}{r} \gamma^r \text{ converges to } N(\mu_i, \sigma_i^2)$$

in distribution for any integer $i \geq 1$, where μ_i and σ_i^2 are constants depending on γ, β and i only.

Lemma 2.2. *Let $\lambda_1, \ldots, \lambda_n$ be random variables with the density as in (2.1). Assume $n\beta/(2a) \to \gamma \in (0, 1]$, and let $\gamma_{\min} = (\sqrt{\gamma} - 1)^2$ and $\gamma_{\max} = (\sqrt{\gamma} + 1)^2$. Let μ_n be the empirical distribution of $Y_i := \lambda_i \gamma/(n\beta)$ for $i = 1, 2, \ldots, n$. Then μ_n converges weakly to the distribution μ_∞ with density $f_\gamma(x)$ as in (1.5) almost surely. Moreover, $\liminf_{n\to\infty} Y_{\max}(n) \geq \gamma_{\max}$ a.s. and $\limsup_{n\to\infty} Y_{\min}(n) \leq \gamma_{\min}$ a.s., where $Y_{\max}(n) = \max\{Y_1, \ldots, Y_n\}$ and $Y_{\min}(n) = \min\{Y_1, \ldots, Y_n\}$.*

Proof. The first part of the above lemma is obtained in [15]. Now we prove the second part. For any integer $k \geq 1$, it is easy to see

$$Y_{\max}(n) \geq \left(\int_0^\infty y^k \, \mathrm{d}\mu_n(y)\right)^{1/k} \quad \text{and} \quad \frac{1}{Y_{\min}(n)} \geq \left(\int_0^\infty y^{-k} \, \mathrm{d}\mu_n(y)\right)^{1/k}.$$

Since μ_n converges weakly to μ_∞ almost surely, by the Fatou lemma,

$$\liminf_{n\to\infty} Y_{\max}(n) \geq \left(\int_{\gamma_{\min}}^{\gamma_{\max}} y^k f_\gamma(y) \, \mathrm{d}y\right)^{1/k} \quad \text{a.s.} \quad \text{and}$$

$$\liminf_{n\to\infty} \frac{1}{Y_{\min}(n)} \geq \left(\int_{\gamma_{\min}}^{\gamma_{\max}} y^{-k} f_\gamma(y) \, \mathrm{d}y\right)^{1/k} \quad \text{a.s.}$$

for any integer $k \geq 1$. Letting $k \to \infty$, we have

$$\liminf_{n\to\infty} Y_{\max}(n) \geq \gamma_{\max} \quad \text{a.s.} \quad \text{and} \quad \limsup_{n\to\infty} Y_{\min}(n) \leq \gamma_{\min} \quad \text{a.s.}$$

□

By following the proof in [41], we have a result below on the β-Laguerre ensembles. It is also reported in Theorem 10.2.2 from [15] without proof.

Lemma 2.3. *Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be random variables with the density function as in (2.1). Set $\lambda_{\max}(n) = \max\{\lambda_1, \ldots, \lambda_n\}$ and $\lambda_{\min} = \min\{\lambda_1, \ldots, \lambda_n\}$. If $n \to \infty, a \to \infty$ and $n\beta/(2a) \to \gamma \in (0, 1]$, then*

$$\frac{\lambda_{\max}(n)}{n} \to \beta\big(1 + \sqrt{\gamma^{-1}}\big)^2 \quad \text{a.s.} \quad \text{and} \quad \frac{\lambda_{\min}(n)}{n} \to \beta\big(1 - \sqrt{\gamma^{-1}}\big)^2 \quad \text{a.s.}$$

as $n \to \infty$.

3. The proof of Theorem 6

First, we give the outline of the proof of Theorem 6. Recalling $f_{\beta,a_1,a_2}(\lambda)$ in (1.1), we see the part $\prod_{i=1}^n (1-\lambda_i)^{a_2-p}$ appeared in the expression of $f_{\beta,a_1,a_2}(\lambda)$. The basic idea is to make $(1-\lambda_i)^{a_2-p} \sim \mathrm{e}^{-(a_2-p)\lambda_i}$ after rescaling λ_i so that it is small enough. Then the beta-Jacobi ensemble is asymptotically the beta-Laguerre ensemble. Concretely, to show that $\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| \to 0$, we write

$$\big\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\big\| = E\big|K_n \cdot L_n(\mu) - 1\big|, \tag{3.1}$$

where K_n is a constant depending on n, a_1, a_2, and $L_n(\mu_1, \ldots, \mu_n)$ is a function of random variables $\mu_1, \ldots, \mu_n$ whose joint density function is the beta-Laguerre ensemble as in (2.1) (in this step we actually make a measure transformation in evaluating the integral). We then analyze the constant K_n very precisely, and study random variable L_n by using the known results on the beta-Laguerre ensemble. Eventually we obtain from the two subtle estimates that $K_n \cdot L_n(\mu) \to 1$ in probability, which concludes (3.1) by a uniform integrability property of $K_n L_n$.

Now we start to prove Theorem 6 by developing several lemmas according to the outline given above.

Lemma 3.1. *Let $n \geq 2$. Let $\mu = (\mu_1, \ldots, \mu_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_n)$ be random variables with density functions $f_{\beta,a_1}(\mu)$ as in* (2.1) (*taking $a = a_1$*) *and $f_{\beta,a_1,a_2}(\lambda)$ in* (1.1), *respectively. Then*

$$\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| = E|K_n \cdot L_n(\mu) - 1|$$

and $E(K_n \cdot L_n(\mu)) = 1$, where the expectation is taken under the density function in (2.1), *and*

$$K_n = a_2^{-na_1} \cdot \prod_{i=0}^{n-1} \frac{\Gamma(a_1 + a_2 - (\beta/2)i)}{\Gamma(a_2 - (\beta/2)i)} \quad \textit{and} \tag{3.2}$$

$$L_n(\mu) = \mathrm{e}^{(1/2)\sum_{i=1}^n \mu_i} \cdot \prod_{i=1}^n \left(1 - \frac{\mu_i}{2a_2}\right)^{a_2 - p} \cdot I\left(\max_{1\leq i\leq n} \mu_i \leq 2a_2\right). \tag{3.3}$$

Proof. It is enough to show

$$\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| = \int_{[0,\infty)^n} |K_n \cdot L_n(\mu) - 1| \cdot f_{\beta,a_1}(\mu)\,\mathrm{d}\mu. \tag{3.4}$$

First, since $p = \frac{\beta}{2}(n-1) + 1$, we have $n(n-1)\beta/2 + n(a_1 - p) + n = na_1$. It is easy to see that the density function of $\theta := 2a_2\lambda$ is

$$\begin{aligned} &g_{\beta,a_1,a_2}(\theta) \\ &\quad := c_J^{\beta,a_1,a_2} \left(\frac{1}{2a_2}\right)^{n(n-1)\beta/2 + n(a_1-p)+n} \prod_{1\leq i<j\leq n} |\theta_i - \theta_j|^\beta \cdot \prod_{i=1}^n \theta_i^{a_1-p} \left(1 - \frac{\theta_i}{2a_2}\right)^{a_2-p} \\ &\quad = c_J^{\beta,a_1,a_2} \left(\frac{1}{2a_2}\right)^{na_1} \prod_{1\leq i<j\leq n} |\theta_i - \theta_j|^\beta \cdot \prod_{i=1}^n \theta_i^{a_1-p} \left(1 - \frac{\theta_i}{2a_2}\right)^{a_2-p} \end{aligned}$$

for $0 \leq \theta_i \leq 2a_2$ and $1 \leq i \leq n$, and is equal to zero, otherwise. Therefore,

$$\begin{aligned} \|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| &= \int_{[0,\infty)^n} |g_{\beta,a_1,a_2}(\mu) - f_{\beta,a_1}(\mu)|\,\mathrm{d}\mu \\ &= \int_{[0,\infty)^n} \left|\frac{g_{\beta,a_1,a_2}(\mu)}{f_{\beta,a_1}(\mu)} - 1\right| \cdot f_{\beta,a_1}(\mu)\,\mathrm{d}\mu. \end{aligned} \tag{3.5}$$

Now, review $f_{\beta,a_1}(\mu)$ as in (2.1) to see that

$$\frac{g_{\beta,a_1,a_2}(\mu)}{f_{\beta,a_1}(\mu)} = \frac{c_J^{\beta,a_1,a_2}}{c_L^{\beta,a_1}} \left(\frac{1}{2a_2}\right)^{na_1} \cdot \prod_{i=1}^n \left(1 - \frac{\mu_i}{2a_2}\right)^{a_2-p} \cdot \mathrm{e}^{\sum_{i=1}^n \mu_i/2}$$

for $0 \le \mu_i \le 2a_2, i = 1, 2, \ldots, n$, and is zero, otherwise. It is easy to check that

$$\frac{c_J^{\beta,a_1,a_2}}{c_L^{\beta,a_1}}\left(\frac{1}{2a_2}\right)^{na_1} = 2^{na_1}\cdot\left(\frac{1}{2a_2}\right)^{na_1}\prod_{j=1}^{n}\frac{\Gamma(a_1+a_2-(\beta/2)(n-j))}{\Gamma(a_2-(\beta/2)(n-j))}$$
$$= a_2^{-na_1}\prod_{i=0}^{n-1}\frac{\Gamma(a_1+a_2-(\beta/2)i)}{\Gamma(a_2-(\beta/2)i)} = K_n.$$

Thus, $g_{\beta,a_1,a_2}(\mu)/f_{\beta,a_1}(\mu) = K_n \cdot L_n(\mu)$, which together with (3.4) and (3.5) yields the first conclusion. Finally,

$$E\big(K_n \cdot L_n(\mu)\big) = \int \frac{g_{\beta,a_1,a_2}(\mu)}{f_{\beta,a_1}(\mu)}\cdot f_{\beta,a_1}(\mu)\,\mathrm{d}\mu = \int g_{\beta,a_1,a_2}(\mu)\,\mathrm{d}\mu = 1.$$ □

Lemma 3.2. *Let $h(x) = x\log x$ for $x > 0$. For a fixed constant $\beta > 0$, an integer $n \ge 1$ and variables $a_1 > 0$ and $a_2 > 0$, set $b_1 = \frac{2}{\beta}a_1$ and $b_2 = \frac{2}{\beta}a_2$. If $n\to\infty$, $a_1\to\infty$ and $a_2\to\infty$ in a way that $a_1 = \mathrm{o}(\sqrt{a_2})$ and $n = \mathrm{o}(\sqrt{a_2})$, then*

$$\sum_{i=1}^{n}\big\{h(b_1+b_2-i+1)-h(b_2-i+1)\big\} = nb_1\left(1+\log b_2+\frac{b_1-n}{2b_2}\right)+\mathrm{o}(1).$$

Proof. Note that $h'(x) = 1+\log x$, $h''(x) = 1/x$ and $h^{(3)}(x) = -1/x^2$. Given $x_0 > 0$, for any $\Delta x > -x_0$, by the Taylor expansion,

$$h(x_0+\Delta x)-h(x_0) = h'(x_0)\Delta x+\frac{1}{2}h''(x_0)(\Delta x)^2+\frac{1}{6}h^{(3)}(\xi)(\Delta x)^3$$
$$= (1+\log x_0)\Delta x+\frac{1}{2x_0}(\Delta x)^2-\frac{1}{6\xi^2}(\Delta x)^3,$$

where ξ is between x_0 and $x_0+\Delta x$. Now take $x_0 = b_2-i+1$ and $\Delta x = b_1$, we have that

$$\begin{aligned}&h(b_1+b_2-i+1)-h(b_2-i+1)\\&\quad= b_1\big(1+\log(b_2-i+1)\big)+\frac{b_1^2}{2}\cdot\frac{1}{b_2-i+1}+\mathrm{O}\left(\frac{b_1^3}{b_2^2}\right)\end{aligned} \tag{3.6}$$

uniformly for all $1 \le i \le n$. Obviously,

$$\log(b_2-i+1) = \log b_2+\log\left(1-\frac{i-1}{b_2}\right) = \log b_2-\frac{i-1}{b_2}+\mathrm{O}\left(\frac{n^2}{b_2^2}\right) \tag{3.7}$$

uniformly over all $1 \le i \le n$ as

$$n\to\infty,\ b_1\to\infty \text{ and } b_2\to\infty \text{ such that}\qquad b_1 = \mathrm{o}(\sqrt{b_2})\quad\text{and}\quad n = \mathrm{o}(\sqrt{b_2}). \tag{3.8}$$

Now,

$$\frac{b_1^2}{2}\cdot\frac{1}{b_2-i+1}=\frac{b_1^2}{2b_2}+\frac{b_1^2}{2}\cdot\left(\frac{1}{b_2-i+1}-\frac{1}{b_2}\right)$$
$$=\frac{b_1^2}{2b_2}+\frac{b_1^2}{2}\cdot\frac{i-1}{b_2(b_2-i+1)}$$
$$=\frac{b_1^2}{2b_2}+\mathrm{O}\left(\frac{nb_1^2}{b_2^2}\right)$$

uniformly for all $1\le i\le n$ as (3.8) holds. Therefore, by (3.6) and (3.7),

$$h(b_1+b_2-i+1)-h(b_2-i+1)$$
$$=b_1+b_1\log b_2-\frac{b_1(i-1)}{b_2}+\frac{b_1^2}{2b_2}+\mathrm{O}\left(\frac{b_1^3+n^2b_1+nb_1^2}{b_2^2}\right)$$

uniformly for all $1\le i\le n$ as (3.8) holds. Thus,

$$\sum_{i=1}^n\bigl\{h(b_1+b_2-i+1)-h(b_2-i+1)\bigr\}$$
$$=nb_1+nb_1\log b_2-\frac{b_1n(n-1)}{2b_2}+\frac{nb_1^2}{2b_2}+n\cdot\mathrm{O}\left(\frac{b_1^3+n^2b_1+nb_1^2}{b_2^2}\right)$$
$$=nb_1\left(1+\log b_2+\frac{b_1-n}{2b_2}\right)+\frac{b_1n}{2b_2}+n\cdot\mathrm{O}\left(\frac{b_1^3+n^2b_1+nb_1^2}{b_2^2}\right).$$

The conclusion follows since the last two terms are all of order $\mathrm{o}(1)$ by (3.8). □

Lemma 3.3. *Let K_n be as in* (3.2). *Assuming* (1.3), *we then have*

$$K_n=\exp\left\{\frac{(1-\gamma)\beta^2n^3}{8a_2\gamma^2}+\mathrm{o}(1)\right\}. \tag{3.9}$$

Proof. We claim that it suffices to prove

$$K_n=\exp\left\{\frac{na_1(a_1-(\beta/2)n)}{2a_2}+\mathrm{o}(1)\right\} \tag{3.10}$$

under assumption (3.8). If this is true, under the condition that $n\beta/(2a_1)\to\gamma\le 1$, it is easy to check that

$$\frac{na_1(a_1-(\beta/2)n)}{2a_2}=\frac{\beta^2n^3t_n(t_n-1)}{8a_2}=\frac{(1-\gamma)\beta^2n^3}{8a_2\gamma^2}+\mathrm{o}(1)$$

as (3.8) holds, where $t_n := 2a_1/(\beta n) \to \gamma^{-1}$. Thus, (3.9) is obtained.

Now we prove (3.10). Set $\alpha = \frac{\beta}{2}$, $b_1 = \frac{2}{\beta}a_1$ and $b_2 = \frac{2}{\beta}a_2$. It is easy to see that

$$K_n = \left(\frac{1}{\alpha b_2}\right)^{n\alpha b_1} \cdot \prod_{i=1}^{n} \frac{\Gamma(\alpha(b_1+b_2-i+1))}{\Gamma(\alpha(b_2-i+1))}.$$

Recall the Stirling formula:

$$\log\Gamma(z) = z\log z - z - \frac{1}{2}\log z + \log\sqrt{2\pi} + \frac{1}{12z} + \mathrm{O}\left(\frac{1}{x^3}\right)$$

as $x = \mathrm{Re}(z) \to +\infty$, where $\Gamma(z) = \int_0^\infty t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t$ with $\mathrm{Re}(z) > 0$, see, for example, page 368 from [27] or (37) on page 204 from [2]. It follows that

$$\begin{aligned}
&\log K_n &(3.11)\\
&\quad = -\alpha n b_1 \log(\alpha b_2)\\
&\qquad + \sum_{i=1}^{n}\bigl\{\alpha(b_1+b_2-i+1)\log\alpha(b_1+b_2-i+1)\\
&\qquad\qquad - \alpha(b_2-i+1)\log\alpha(b_2-i+1)\bigr\} &(3.12)\\
&\qquad - \alpha n b_1 - \frac{1}{2}\sum_{i=1}^{n}\log\frac{b_1+b_2-i+1}{b_2-i+1} + \mathrm{O}\left(\frac{1}{b_2-n}\right) &(3.13)
\end{aligned}$$

as (3.8) holds.

Now, write $(\alpha x)\log(\alpha x) = (\alpha\log\alpha)x + \alpha(x\log x)$ and set $h(x) = x\log x$ for $x > 0$. Calculating the difference between the two terms in the sum of (3.12), we know that the whole sum in (3.12) is identical to

$$\begin{aligned}
&\alpha(\log\alpha)nb_1 + \alpha\sum_{i=1}^{n}\bigl(h(b_1+b_2-i+1) - h(b_2-i+1)\bigr)\\
&\quad = \alpha(\log\alpha)nb_1 + \alpha n b_1\left(1 + \log b_2 + \frac{b_1-n}{2b_2}\right) + \mathrm{o}(1) \qquad (3.14)\\
&\quad = \alpha n b_1 + (\alpha n b_1)\log(\alpha b_2) + \alpha n b_1 \cdot \frac{b_1-n}{2b_2} + \mathrm{o}(1)
\end{aligned}$$

by Lemma 3.2. From the fact that $\log(1+x) \le x$ for any $x \ge 0$, we have

$$0 < \sum_{i=1}^{n}\log\frac{b_1+b_2-i+1}{b_2-i+1} = \sum_{i=1}^{n}\log\left(1 + \frac{b_1}{b_2-i+1}\right) \le \frac{nb_1}{b_2-n}$$

for any $b_2 > n$. Thus, the sum of the three terms in (3.13) is equal to $-\alpha n b_1 + O(\frac{nb_1}{b_2})$ as (3.8) holds. Combining this and (3.11)–(3.14), we get

$$\log K_n = \alpha n b_1 \cdot \frac{b_1 - n}{2b_2} + o(1) = \frac{na_1(a_1 - (\beta/2)n)}{2a_2} + o(1)$$

as (3.8) holds. This gives (3.10). □

In the proofs next, we use $o_P(1)$ to denote a random variable that approaches zero in probability.

Lemma 3.4. *Let $\mu = (\mu_1, \ldots, \mu_n)$ be a random variable with the density function as in (2.1) with $a = a_1$. Let $L_n(\mu)$ be as in (3.3). If (1.3) holds, then*

$$\exp\left\{\frac{(1-\gamma)\beta^2 n^3}{8a_2\gamma^2}\right\} \cdot L_n(\mu) \to 1$$

in probability as $n \to \infty$.

Proof. From (3.3), we see that

$$L_n(\mu) = e^{(1/2)\sum_{i=1}^n \mu_i} \cdot \prod_{i=1}^n \left(1 - \frac{\mu_i}{2a_2}\right)^{a_2 - p} \cdot I\left(\max_{1 \le i \le n} \mu_i \le 2a_2\right).$$

By Lemma 2.3, since $\frac{n\beta}{2a_1} \to \gamma \in (0, 1]$ by (1.3),

$$\frac{\max_{1 \le i \le n} \mu_i}{n} \to \beta\left(1 + \sqrt{\gamma^{-1}}\right)^2 \tag{3.15}$$

in probability as $n \to \infty$. Since $n = o(\sqrt{a_2})$, choose $\delta_n = (n\sqrt{a_2})^{1/2}$, then $\delta_n/n \to \infty$ and $\delta_n/\sqrt{a_2} \to 0$ as taking the limit as in (1.3). Therefore, to prove the lemma, it is enough to show

$$\exp\left\{\frac{(1-\gamma)\beta^2 n^3}{8a_2\gamma^2}\right\} \cdot \tilde{L}_n(\mu) \to 1 \tag{3.16}$$

in probability as $n \to \infty$, where

$$\tilde{L}_n(\mu) := e^{(1/2)\sum_{i=1}^n \mu_i} \cdot \prod_{i=1}^n \left(1 - \frac{\mu_i}{2a_2}\right)^{a_2 - p} \cdot I\left(\max_{1 \le i \le n} \mu_i \le \delta_n\right). \tag{3.17}$$

This is because, for any two sequences random variables $\{\xi_n; n \ge 1\}$ and $\{\eta_n; n \ge 1\}$, if $\xi_n \to 1$ in probability and $P(\xi_n \ne \eta_n) \to 0$ as $n \to \infty$, then $\eta_n \to 1$ in probability as $n \to \infty$. Rewrite

$$\tilde{L}_n(\mu) = \exp\left\{\frac{1}{2}\sum_{i=1}^n \mu_i + (a_2 - p)\sum_{i=1}^n \log\left(1 - \frac{\mu_i}{2a_2}\right)\right\}$$

on $\Omega_n := \{\max_{1\le i\le n} \mu_i \le \delta_n\}$. Noticing $\log(1-x) = -x - (x^2/2) + O(x^3)$ as $x \to 0$,

$$\sum_{i=1}^{n} \log\left(1 - \frac{\mu_i}{2a_2}\right) = -\frac{1}{2a_2}\sum_{i=1}^{n}\mu_i - \frac{1}{8a_2^2}\sum_{i=1}^{n}\mu_i^2 + O\left(\frac{1}{a_2^3}\sum_{i=1}^{n}\mu_i^3\right) \tag{3.18}$$

on Ω_n. Now, on Ω_n again,

$$\frac{1}{a_2^3}\sum_{i=1}^{n}\mu_i^3 \le \frac{n(\delta_n)^3}{a_2^3} = \left(\frac{\delta_n}{\sqrt{a_2}}\right)^3 \cdot \frac{n}{\sqrt{a_2}} \cdot \frac{1}{a_2} \to 0 \tag{3.19}$$

as taking the limit in (1.3). Recall $p = 1 + \frac{\beta}{2}(n-1)$. We have from (3.18) and (3.19) that, on Ω_n,

$$\begin{aligned}
&(a_2 - p)\sum_{i=1}^{n}\log\left(1 - \frac{\mu_i}{2a_2}\right) \\
&\quad = -\frac{a_2 - p}{2a_2}\left(-\frac{\beta n^2}{\gamma} + \sum_{i=1}^{n}\mu_i\right) - \frac{(a_2 - p)\beta n^2}{2a_2\gamma} \\
&\qquad - \frac{a_2 - p}{8a_2^2}\left(-\frac{\beta^2 n^3}{\gamma^2}(1+\gamma) + \sum_{i=1}^{n}\mu_i^2\right) - \frac{\beta^2 n^3 (a_2 - p)}{8a_2^2\gamma^2}(1+\gamma) + O\left(\left(\frac{\delta_n}{\sqrt{a_2}}\right)^3 \cdot \frac{n}{\sqrt{a_2}}\right)
\end{aligned}$$

by (1.3). From Lemma 2.1, as $n \to \infty$,

$$\frac{1}{n^2}\sum_{i=1}^{n}\mu_i \xrightarrow{P} \frac{\beta}{\gamma}; \qquad \frac{1}{n}\sum_{i=1}^{n}\mu_i - \frac{\beta}{\gamma}n \Rightarrow N(0, \sigma_1^2); \tag{3.20}$$

$$\begin{aligned}
&\frac{1}{n^3}\sum_{i=1}^{n}\mu_i^2 \xrightarrow{P} \frac{\beta^2}{\gamma^2}\left(1 + \frac{1}{2}\cdot 2\gamma\right) = \frac{\beta^2}{\gamma^2}(1+\gamma); \\
&\frac{1}{n^2}\sum_{i=1}^{n}\mu_i^2 - \frac{\beta^2}{\gamma^2}(1+\gamma)n \Rightarrow N(0, \sigma_2^2),
\end{aligned} \tag{3.21}$$

where σ_1, σ_2 are constants depending on γ only, the notation "$\xrightarrow{P}$" means "converges in probability to" and "$\Rightarrow$" means "converges weakly to". Now, write $(a_2 - p)/2a_2 = (1/2) - p/2a_2$, then

$$\begin{aligned}
&-\frac{a_2 - p}{2a_2}\left(-\frac{\beta n^2}{\gamma} + \sum_{i=1}^{n}\mu_i\right) \\
&\quad = \frac{\beta n^2}{2\gamma} - \frac{1}{2}\sum_{i=1}^{n}\mu_i + \frac{pn}{2a_2}\cdot\frac{1}{n}\left(-\frac{\beta n^2}{\gamma} + \sum_{i=1}^{n}\mu_i\right) \\
&\quad = \frac{\beta n^2}{2\gamma} - \frac{1}{2}\sum_{i=1}^{n}\mu_i + o_P(1)
\end{aligned}$$

by (3.20) and (1.3). Also, under the same condition, $(a_2 - p)n^2/a_2^2 = \mathrm{O}(n^2/a_2) = \mathrm{o}(1)$. It follows from (3.21) that

$$-\frac{a_2 - p}{8a_2^2}\left(-\frac{\beta^2 n^3}{\gamma^2}(1+\gamma) + \sum_{i=1}^n \mu_i^2\right) = -\frac{(a_2 - p)n^2}{8a_2^2}\left(-\frac{\beta^2 n}{\gamma^2}(1+\gamma) + \frac{1}{n^2}\sum_{i=1}^n \mu_i^2\right) \to 0$$

in probability as taking the limit in (1.3). In summary, combining all the computations above,

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^n \mu_i + (a_2 - p)\sum_{i=1}^n \log\left(1 - \frac{\mu_i}{2a_2}\right) \\
&\quad = \frac{\beta n^2}{2\gamma} - \frac{(a_2 - p)\beta n^2}{2a_2\gamma} - \frac{\beta^2 n^3 (a_2 - p)}{8a_2^2\gamma^2}(1+\gamma) + \mathrm{o}_P(1) \\
&\quad = \frac{\beta p n^2}{2a_2\gamma} - \frac{\beta^2 n^3}{8a_2\gamma^2}(1+\gamma) + \frac{\beta^2 n^3 p}{8a_2^2\gamma^2}(1+\gamma) + \mathrm{o}_P(1)
\end{aligned}$$

on Ω_n. Now, since $p = 1 + \frac{\beta}{2}(n-1)$, $n/\sqrt{a_2} \to 0$, we see that

$$\frac{\beta p n^2}{2a_2\gamma} = \frac{\beta^2 n^3}{4a_2\gamma} + \mathrm{o}(1) \quad \text{and} \quad \frac{\beta^2 n^3 p}{8a_2^2\gamma^2}(1+\gamma) \to 0$$

by (1.3). Thus, on Ω_n,

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^n \mu_i + (a_2 - p)\sum_{i=1}^n \log\left(1 - \frac{\mu_i}{2a_2}\right) \\
&\quad = \frac{\beta^2 n^3}{4a_2\gamma} - \frac{\beta^2 n^3}{8a_2\gamma^2}(1+\gamma) + \mathrm{o}_P(1) = \frac{(\gamma - 1)\beta^2 n^3}{8a_2\gamma^2} + \mathrm{o}_P(1)
\end{aligned}$$

as taking the limit in (1.3). By reviewing (3.17), we conclude (3.16). □

The following is a variant of the Scheffé Lemma, see, for example, Corollary 4.2.4 from [6].

Lemma 3.5. *Let $\{X_n; n \geq 1\}$ be a sequence of non-negative random variables. If $X_n \to 1$ in probability and $EX_n \to 1$ as $n \to \infty$, then $E|X_n - 1| \to 0$ as $n \to \infty$.*

Proof of Theorem 6. It is known from Lemma 3.1 that

$$\left\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\right\| = E\left|K_n \cdot L_n(\mu) - 1\right|$$

with $E(K_n \cdot L_n(\mu)) = 1$ for all $n \geq 2$, where μ has density $f_{\beta,a_1}(\mu)$ as in (2.1). By Lemmas 3.3 and 3.4,

$$K_n = \exp\left\{\frac{(1-\gamma)\beta^2 n^3}{8a_2\gamma^2} + \mathrm{o}(1)\right\} \quad \text{and} \quad \exp\left\{\frac{(1-\gamma)\beta^2 n^3}{8a_2\gamma^2}\right\} \cdot L_n(\mu) \to 1$$

in probability as taking the limit in (1.3). These imply that $K_n \cdot L_n(\mu) \to 1$ in probability as taking the same limit. Then the desired conclusion follows from Lemma 3.5. □

4. The proofs of Theorems 1–5

After proved Theorem 6 in Section 3, we now are ready to use it to prove Theorems 1–5 stated in Section 1.

Proof of Theorem 1. Set

$$\nu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{(a_2/n)\lambda_i'}$$

for $(\lambda_1', \ldots, \lambda_n') \in [0, +\infty)^n$. Then, recall the definition of d in (1.4), by the triangle inequality,

$$\begin{aligned} \left|d(\mu_n, \mu_0) - d(\nu_n, \mu_0)\right| &\leq d(\mu_n, \nu_n) \\ &= \sup_{\|f\|_{\mathrm{BL}} \leq 1} \left|\frac{1}{n}\sum_{i=1}^{n}\left(f\left(n^{-1}a_2\lambda_i\right) - f\left(n^{-1}a_2\lambda_i'\right)\right)\right| \\ &\leq \frac{a_2}{n} \cdot \max_{1\leq i\leq n}\left|\lambda_i - \lambda_i'\right|, \end{aligned}$$

where the Lipschitz inequality $|f(x) - f(y)| \leq |x - y|$ is used in the last step. This says that $d(\mu_n, \mu_0)$, as a function of $(\lambda_1, \ldots, \lambda_n)$, is continuous for each $n \geq 2$. Thus, for any $\varepsilon > 0$, there exists a (non-random) Borel set $A \subset \mathbb{R}^n$ such that $\{d(\mu_n, \mu_0) \geq \varepsilon\} = \{(\lambda_1, \ldots, \lambda_n) \in A\}$. Then, by the definition of the variation norm in (1.7) we see that

$$P\left(d(\mu_n, \mu_0) \geq \varepsilon\right) \leq P\left(d\left(\mu_n', \mu_0\right) \geq \varepsilon\right) + \left\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\right\| \tag{4.1}$$

for any $\varepsilon > 0$, where $\mu_n' = (1/n)\sum_{i=1}^{n} \delta_{\mu_i/(2n)}$ and $\mu = (\mu_1, \ldots, \mu_n)$ has density $f_{\beta,a_1}(\mu)$ as in (2.1) with $a = a_1$ and $n\beta/2a_1 \to \gamma \in (0, 1]$. By Lemma 2.2, with probability one,

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\mu_i\gamma/(n\beta)} \qquad \text{converges weakly to } \mu_\infty \tag{4.2}$$

with density $f_\gamma(x)$ as in (1.5). Write $\mu_i/(2n) = (\mu_i\gamma/n\beta)c^{-1}$, where $c = 2\gamma/\beta$. Then, by (4.2), with probability one, μ_n' converges weakly to μ_0, where μ_0 has density function $c \cdot f_\gamma(cx)$. Equivalently, $d(\mu_n', \mu_0) \to 0$ almost surely. This, (4.1) and Theorem 6 prove the theorem. □

Proof of Theorem 2. First, $\lambda_{\max}(n)$ and $\lambda_{\min}(n)$ are continuous functions of $\lambda_1, \dots, \lambda_n$ for any $n \ge 1$. Then

$$\begin{aligned} &P\left(\left|\frac{a_2}{n}\lambda_{\max}(n) - \frac{\beta(1+\sqrt{\gamma})^2}{2\gamma}\right| \ge \varepsilon\right) \\ &\quad \le P\left(\left|\frac{1}{2n}\mu_{\max}(n) - \frac{\beta(1+\sqrt{\gamma})^2}{2\gamma}\right| \ge \varepsilon\right) + \left\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\right\| \end{aligned} \tag{4.3}$$

for any $\varepsilon > 0$, where $\mu = (\mu_1, \dots, \mu_n)$ has density $f_{\beta,a_1}(\mu)$ as in (2.1) with $a = a_1$ and $n\beta/2a_1 \to \gamma \in (0,1]$. From Lemma 2.3, we know $\mu_{\max}(n)/(2n) \to \beta(1+\sqrt{\gamma^{-1}})^2/2 = \beta(1+\sqrt{\gamma})^2/(2\gamma)$ in probability. This together with (4.3) and Theorem 6 yields the desired conclusion. By the same argument, $(a_2/n)\lambda_{\min}(n)$ converges to $\beta(1-\sqrt{\gamma})^2/(2\gamma)$ in probability. □

Proof of Theorem 3. Evidently,

$$\left|P\big((X_1, \dots, X_k) \in A\big) - P\big((Y_1, \dots, Y_k) \in A\big)\right| \le \left\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\right\|$$

for any Borel set $A \in \mathbb{R}^k$, where

$$\begin{aligned} Y_i &= \sum_{j=1}^{n}\left(\frac{c}{2n}\mu_j\right)^i - n\sum_{r=0}^{i-1}\frac{1}{r+1}\binom{i}{r}\binom{i-1}{r}\gamma^r \\ &= \sum_{j=1}^{n}\left(\frac{\gamma}{n\beta}\mu_j\right)^i - n\sum_{r=0}^{i-1}\frac{1}{r+1}\binom{i}{r}\binom{i-1}{r}\gamma^r \end{aligned}$$

for $i \ge 1$ (since $c = 2\gamma/\beta$), and $\mu = (\mu_1, \dots, \mu_n)$ has density $f_{\beta,a_1}(\mu)$ as in (2.1) with $a = a_1$ and $n\beta/2a_1 \to \gamma \in (0,1]$. The conclusion then follows from this, Theorem 1.5 in [17] and Theorem 6. □

Proof of Theorem 4. The assumption that $2\beta^{-1}a_1 - n = c$ and $n = o(\sqrt{a_2})$ imply that $n\beta/2a_1 \to 1$ and $a_1 = o(\sqrt{a_2})$. Thus, Theorem 6 holds.

Let $(\theta_1, \dots, \theta_n)$ have density f_{β,a_1} as in (2.1) with $a = a_1$. Noticing, "$\beta\lambda_i$" and "a" in Theorem 1 from [40] correspond to "θ_i" and "c" here, respectively. By Theorem 1 from [40], for fixed integer $k \ge 1$,

$$\left(\frac{n}{\beta}\theta^{(n)}, \dots, \frac{n}{\beta}\theta^{(n-k+1)}\right) \qquad \text{converges weakly to } \big(\Lambda_0(\beta, c), \dots, \Lambda_{k-1}(\beta, c)\big)$$

as $n \to \infty$. By Theorem 6,

$$P\big((2a_2\lambda_1, \dots, 2a_2\lambda_n) \in B_n\big) - P\big((\theta_1, \dots, \theta_n) \in B_n\big) \to 0$$

for any Borel set $B_n \subset \mathbb{R}^n$ for $n \ge 1$. From the Weyl perturbation theorem, we know that $\lambda^{(i)}$ is a continuous function of $(\lambda_1, \dots, \lambda_n)$ for any $1 \le i \le n$. Combining the above two limits, we obtain

$$\left(\frac{2a_2n}{\beta}\lambda^{(n)},\ldots,\frac{2a_2n}{\beta}\lambda^{(n-k+1)}\right) \quad \text{converges weakly to } \big(\Lambda_0(\beta,c),\ldots,\Lambda_{k-1}(\beta,c)\big)$$

as $n\to\infty$ and $a_2\to\infty$ with $n=o(\sqrt{a_2})$. The proof is complete. □

Proof of Theorem 5. Recalling (2.1), let

$$\tilde{m}_n=\big(\sqrt{n}+\sqrt{2\beta^{-1}a}\big)^2 \quad\text{and}\quad \tilde{\sigma}_n=\frac{(2\beta^{-1}na)^{1/6}}{(\sqrt{n}+\sqrt{2\beta^{-1}a})^{4/3}}. \tag{4.4}$$

Let $(\theta_1,\ldots,\theta_n)$ have density $f_{\beta,a}$ as in (2.1). Noticing, "$\beta\lambda_i$" in Theorem 1.4 from [39] corresponds to "θ_i" here; "κ" in Theorem 1.4 from [39] is equal to $2\beta^{-1}a$, and $\frac{\beta}{2}(n-1)+1=p$, and the kth lowest eigenvalue of $\mathcal{H}_\beta$ is the $(n-k+1)$th largest eigenvalue of $-\mathcal{H}_\beta$. Then by Theorem 1.4 from [39],

$$\tilde{\sigma}_n\left(\frac{\theta^{(l)}}{\beta}-\tilde{m}_n\right)_{l=1,\ldots,k} \quad\text{converges weakly to } (\Lambda_1,\ldots,\Lambda_k) \tag{4.5}$$

as $n\to\infty$ and $a\to\infty$ such that n/a converges to a non-zero, finite constant. In other words,

$$P\left(\tilde{\sigma}_n\left(\frac{\theta^{(l)}}{\beta}-\tilde{m}_n\right)_{l=1,\ldots,k}\in A\right)\to P\big((\Lambda_1,\ldots,\Lambda_k)\in A\big)$$

for any Borel set $A\subset\mathbb{R}^k$. By Theorem 6, assuming (1.3) and $a=a_1$,

$$P\big((2a_2\lambda_1,\ldots,2a_2\lambda_n)\in B_n\big)-P\big((\theta_1,\ldots,\theta_n)\in B_n\big)\to 0$$

for any Borel set $B_n\subset\mathbb{R}^n$ for $n\geq 1$. The Weyl perturbation theorem says that $g(x):=x^{(l)}$, the lth largest one in $\{x_1,\ldots,x_m\}$, is a continuous function of $(x_1,\ldots,x_m)\in\mathbb{R}^m$ for any $1\leq l\leq m$. Replacing a by a_1 in (4.4), the above two assertions conclude that

$$\sigma_n\left(\frac{2a_2\lambda^{(l)}}{\beta}-m_n\right)_{l=1,\ldots,k} \quad\text{converges weakly to } (\Lambda_1,\ldots,\Lambda_k).$$

□

Acknowledgements

The author thanks Professor Ioana Dumitriu for helpful discussions on the beta ensembles. The author also thanks two anonymous referees very much for their very careful reading and many constructive suggestions. Supported in part by NSF#DMS-0449365.

References

[1] Absil, P.A., Edelman, A. and Koev, P. (2006). On the largest principal angle between random subspaces. *Linear Algebra Appl*. **414** 288–294. MR2209246

[2] Ahlfors, L.V. (1979). *Complex Analysis*, 3rd ed. New York: McGraw-Hill Book Co. MR0510197

[3] Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0771294

[4] Aomoto, K. (1987). Jacobi polynomials associated with Selberg integrals. *SIAM J*. *Math*. *Anal*. **18** 545–549. MR0876291

[5] Baker, T.H. and Forrester, P.J. (1997). The Calogero–Sutherland model and generalized classical polynomials. *Comm*. *Math*. *Phys*. **188** 175–216. MR1471336

[6] Chow, Y.S. and Teicher, H. (1997). *Probability Theory*: *Independence*, *Interchangeability*, *Martingales*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR1476912

[7] Collins, B. (2003). Intégrales Matricielles et Probabilitiés Non-commutatives. Thèse de Doctorat of Univ. Paris 6.

[8] Collins, B. (2005). Product of random projections, Jacobi ensembles and universality problems arising from free probability. *Probab. Theory Related Fields* **133** 315–344. MR2198015

[9] Constantine, A.G. (1963). Some non-central distribution problems in multivariate analysis. *Ann*. *Math*. *Statist*. **34** 1270–1285. MR0181056

[10] Demni, N. (2010). β-Jacobi processes. *Adv. Pure Appl*. *Math*. **1** 325–344. MR2719370

[11] Dette, H. and Nagel, J. (2009). Some asymptotic properties of the spectrum of the Jacobi ensemble. *SIAM J*. *Math*. *Anal*. **41** 1491–1507. MR2556572

[12] Diaconis, P.W., Eaton, M.L. and Lauritzen, S.L. (1992). Finite de Finetti theorems in linear models and multivariate analysis. *Scand*. *J*. *Stat*. **19** 289–315. MR1211786

[13] Dong, Z., Jiang, T. and Li, D. (2012). Circular law and arc law for truncation of random unitary matrix. *J*. *Math*. *Phys*. **53** 013301, 14. MR2919538

[14] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. MR1932358

[15] Dumitriu, I. (2003). Eigenvalue statistics for Beta-ensembles. Ph.D. thesis, Massachusetts Institute of Technology, ProQuest LLC, Ann Arbor, MI. MR2717094

[16] Dumitriu, I. and Edelman, A. (2002). Matrix models for beta ensembles. *J*. *Math*. *Phys*. **43** 5830–5847. MR1936554

[17] Dumitriu, I. and Edelman, A. (2006). Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models. *J*. *Math*. *Phys*. **47** 063302, 36. MR2239975

[18] Dumitriu, I. and Koev, P. (2008). Distributions of the extreme eigenvalues of beta-Jacobi random matrices. *SIAM J*. *Matrix Anal*. *Appl*. **30** 1–6. MR2399565

[19] Dyson, F.J. (1962). Statistical theory of the energy levels of complex systems. I. *J*. *Math*. *Phys*. **3** 140–156. MR0143556

[20] Eaton, M.L. (1983). *Multivariate Statistics*: *A Vector Space Approach*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0716321

[21] Edelman, A. and Rao, N.R. (2005). Random matrix theory. *Acta Numer.* **14** 233–297. MR2168344

[22] Edelman, A. and Sutton, B.D. (2008). The beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems. *Found*. *Comput*. *Math*. **8** 259–285. MR2407033

[23] Edelman, A.S. (1989). Eigenvalues and condition numbers of random matrices. Ph.D. thesis, Massachusetts Institute of Technology, ProQuest LLC, Ann Arbor, MI. MR2941174

[24] Forrester, P.J. (2006). Quantum conductance problems and the Jacobi ensemble. *J*. *Phys*. *A* **39** 6861–6870. MR2233258

[25] Forrester, P.J. (2010). *Log-gases and Random Matrices*. *London Mathematical Society Monographs Series* **34**. Princeton, NJ: Princeton Univ. Press. MR2641363
[26] Forrester, P.J. and Warnaar, S.O. (2008). The importance of the Selberg integral. *Bull*. *Amer. Math*. *Soc*. (*N.S.*) **45** 489–534. MR2434345
[27] Gamelin, T.W. (2001). *Complex Analysis*. *Undergraduate Texts in Mathematics*. New York: Springer. MR1830078
[28] James, A.T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann*. *Math*. *Statist*. **35** 475–501. MR0181057
[29] Jiang, T. (2006). How many entries of a typical orthogonal matrix can be approximated by independent normals? *Ann*. *Probab*. **34** 1497–1529. MR2257653
[30] Jiang, T. (2009). Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles. *Probab. Theory Related Fields* **144** 221–246. MR2480790
[31] Johnstone, I.M. (2008). Multivariate analysis and Jacobi ensembles: Largest eigenvalue, Tracy–Widom limits and rates of convergence. *Ann*. *Statist*. **36** 2638–2716. MR2485010
[32] Kadell, K.W.J. (1997). The Selberg–Jack symmetric functions. *Adv*. *Math*. **130** 33–102. MR1467311
[33] Killip, R. (2008). Gaussian fluctuations for β ensembles. *Int. Math. Res. Not. IMRN* **2008** Art. ID rnn007. MR2428142
[34] Killip, R. and Nenciu, I. (2004). Matrix models for circular ensembles. *Int*. *Math*. *Res*. *Not*. **50** 2665–2701. MR2127367
[35] Lippert, R.A. (2003). A matrix model for the β-Jacobi ensemble. *J. Math*. *Phys*. **44** 4807–4816. MR2008950
[36] Macdonald, I. (1998). *Symmertic Functions and Hall Polynomials*, 2nd ed. *Oxford Mathematical Monographs*. London: Oxford Univ. Press.
[37] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. New York: Wiley. MR0652932
[38] Okounkov, A. and Olshanski, G. (1997). Shifted Jack polynomials, binomial formula, and applications. *Math*. *Res*. *Lett*. **4** 69–78. MR1432811
[39] Ramírez, J., Rider, B. and Virág, B. (2011). Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.* **24** 919–944. MR2813333
[40] Ramírez, J.A. and Rider, B. (2009). Diffusion at the random matrix hard edge. *Comm*. *Math*. *Phys*. **288** 887–906. MR2504858
[41] Silverstein, J.W. (1985). The smallest eigenvalue of a large-dimensional Wishart matrix. *Ann*. *Probab*. **13** 1364–1368. MR0806232
[42] Stanley, R.P. (1989). Some combinatorial properties of Jack symmetric functions. *Adv*. *Math*. **77** 76–115. MR1014073
[43] Wang, M.C. and Uhlenbeck, G.E. (1945). On the theory of the Brownian motion. II. *Rev. Modern Phys*. **17** 323–342. MR0013266

Received November 2010 and revised November 2011