

Limit Theorems for Beta-Jacobi Ensembles

TIEFENG JIANG ¹

Abstract For a large beta-Jacobi ensemble determined by several parameters, under certain restrictions among them, we obtain both the bulk and the edge scaling limits. 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.

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, has density function

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

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

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

The ensemble is associated with 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 [14] and [56] for $\beta = 1, 2$. Other references about the connections between the Jacobi ensembles and statistics are [6, 12, 13, 14, 19, 28, 34, 42, 43, 56].

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 changes. 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, e.g., [9, 27, 32, 67].

¹Supported in part by NSF#DMS-0449365, School of Statistics, University of Minnesota, 224 Church Street, MN55455, tjiang@stat.umn.edu.

Key Words: random matrix, Jacobi ensemble, Laguerre ensemble, beta-ensemble, largest eigenvalue, smallest eigenvalue, empirical distribution, random operator, limiting distribution.

AMS (2000) subject classifications: 60B10, 60B20, 60F05, 60F15, 62H10.

The beta-Jacobi ensembles also have connections to other subjects in mathematics and physics, for instance, lattice gas theory [32, 34], Selberg integrals [33, 54, 57] and Jack functions [4, 50, 61].

Now we briefly recall some research on the beta-Jacobi ensembles. Lippert [52] gives a model to generate the beta-Jacobi ensembles (see also [51] 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 [29] study CS decomposition and singular values about these models. Dumitriu and Koev [22] derive the exact distributions of the largest eigenvalues for the ensembles. Jiang [42] derives the bulk and the edge scaling limits for the beta-Jacobi ensembles for $\beta = 1, 2$ when p and a_1 in (1.1) are of small orders of a_2 . Johnstone [49] obtains the asymptotic distribution of the largest eigenvalues for $\beta = 1, 2$ when a_1, a_2 and p in (1.1) are proportional to each other. Recently, Demni [16] 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, $n \geq 2$ be an integer, a_1 and a_2 be positive variables. The following condition will be used several times later.

$$n \rightarrow \infty, a_1 \rightarrow \infty \text{ and } a_2 \rightarrow \infty \text{ such that } a_1 = o(\sqrt{a_2}), n = o(\sqrt{a_2}) \text{ and } \frac{n\beta}{2a_1} \rightarrow \gamma \in (0, 1]. \quad (1.3)$$

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

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

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

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

Then, for a sequence of probability measures $\{\mu_n; n = 0, 1, 2, \dots\}$ defined on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, we know μ_n converges weakly to μ_0 if and only if $d(\mu_n, \mu_0) \rightarrow 0$ as $n \rightarrow \infty$, see, e.g., [21]. Similarly, we say that a sequence of random variables $\{Z_n; n \geq 1\}$ taking values in \mathbb{R}^k converges weakly (or in distribution) to a Borel probability measure μ on \mathbb{R}^k if $Ef(Z_n) \rightarrow \int_{\mathbb{R}^k} f(x) \mu(dx)$ for any bounded and continuous function $f(x)$ defined on \mathbb{R}^k . This is also equivalent to that $d(\mathcal{L}(Z_n), \mu) \rightarrow 0$ as $n \rightarrow \infty$, where $\mathcal{L}(Z_n)$ is the probability distribution of Z_n , see also [21].

For $\gamma > 0$, 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

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

The following is about the limiting distribution of the empirical eigenvalues of the beta-Jacobi ensembles.

THEOREM 1 Let $\lambda_1, \dots, \lambda_n$ be random variables with density function $f_{\beta, a_1, a_2}(\lambda)$ as in (1.1). Set

$$\mu_n = \frac{1}{n} \sum_{i=1}^n I_{\frac{a_2}{n}} \lambda_i$$

for $n \geq 2$. 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 weak laws of large numbers of the largest and smallest eigenvalues for the beta-Jacobi ensembles.

THEOREM 2 Let $\lambda_1, \dots, \lambda_n$ be random variables with density function $f_{\beta, a_1, a_2}(\lambda)$ as in (1.1). Set $\lambda_{max}(n) = \max\{\lambda_1, \dots, \lambda_n\}$, and $\lambda_{min}(n) = \min\{\lambda_1, \dots, \lambda_n\}$. Assuming (1.3), we have that

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

in probability.

Here is the central limit theorem for the eigenvalues for the model in (1.1).

THEOREM 3 Let $\lambda_1, \dots, \lambda_n$ be random variables with density function $f_{\beta, a_1, a_2}(\lambda)$ as in (1.1). Given integer $k \geq 1$, define

$$X_i = \sum_{j=1}^n \left(\frac{c a_2}{n} \lambda_j \right)^i - n \sum_{r=0}^{i-1} \frac{1}{r+1} \binom{i}{r} \binom{i-1}{r} \gamma^r$$

for $i \geq 1$, where $c = 2\gamma/\beta$ and γ is as in (1.3). Assuming (1.3), then (X_1, \dots, X_k) converges weakly to a multivariate normal distribution $N_k(\mu, \Sigma)$ for some μ and Σ given in Theorem 1.5 from [23].

Thanks to the recent results of Ramírez and Rider [59] and Ramírez, Rider and Virág [58], we are able to investigate the asymptotic distributions of the smallest and largest eigenvalues for the beta-Jacobi ensembles next. Look at the operator

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

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 $0 < \Lambda_0(\beta, a) < \Lambda_1(\beta, a) < \dots \uparrow \infty$ as stated in Theorem 1 from [59].

For a sequence of pairwise different numbers a_1, \dots, a_n , let $a^{(1)} > a^{(2)} > \dots > a^{(n)}$ be their order statistic. The following is the limiting distribution of the first k smallest eigenvalues.

THEOREM 4 Let $\lambda_1, \dots, \lambda_n$ 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 $n \rightarrow \infty$ and $a_2 \rightarrow \infty$ such that $n = o(\sqrt{a_2})$, then $(2\beta^{-1}na_2) \cdot (\lambda^{(n)}, \dots, \lambda^{(n-k+1)})$ converges weakly to $(\Lambda_0(\beta, c), \Lambda_1(\beta, c), \dots, \Lambda_{k-1}(\beta, c))$.

Now look at another random operator

$$-\mathcal{H}_\beta = \frac{d^2}{dx^2} - x - \frac{2}{\sqrt{\beta}} b'_x \quad (1.6)$$

where b_x is a standard Brownian motion on $[0, +\infty)$. For $\lambda \in \mathbb{R}$ and function $\psi(x)$ defined on $[0, +\infty)$ with $\psi(0) = 0$ and $\int_0^\infty ((\psi')^2 + (1+x)\psi^2) dx < \infty$, we say (ψ, λ) is an eigenfunction/eigenvalue pair for $-\mathcal{H}_\beta$ if $\int_0^\infty \psi^2(x) dx = 1$ and

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

holds in the following integration-by-parts sense,

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

Theorem 1.1 from [59] says that, with probability one, for each $k \geq 1$, the set of eigenvalues of $-\mathcal{H}_\beta$ has well-defined k -largest eigenvalues Λ_k . Recall (1.1), set

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

Our last result is about the limiting distribution of the first k largest eigenvalues of the beta-Jacobi ensembles.

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

REMARK 1.1 *In [42], Jiang study Theorems 1, 2, 3 and 5 for $\beta = 1$ and 2, which are special cases of the current theorems. The method used in [42] is the approximation of the entries of Haar-invariant orthogonal or unitary matrices by independent and identically distributed real or complex Gaussian random variables.*

Now, let us state the methodology used in our proofs. In fact, we employ a different approach than the standard ones in the random matrix theory. Some of the standard tools are the moment method in [7, 18, 20, 48, 68], the Stieltjes transformations in [53, 55], the analysis to study the probability density functions of eigenvalues in [2, 8, 46, 47, 64, 65, 66], the large deviation method for obtaining the law of large numbers in [5, 36, 37], and the application of the free probability theory in [10, 11]. A technique on refined estimates of the smallest eigenvalues is used in [63].

Another way to study random matrices is using the known conclusions, and connect them with the target ones by approximation. For example, large sample correlation matrices can be approximated by the Wishart matrices [45]; a large dimensional, Haar-invariant matrix from the classical compact groups can be approximated by matrices with independent Gaussian entries [40, 44]. The study of local statistics of the Wigner matrices with non-Gaussian entries can be approximated by Gaussian entries [62].

In this paper, we approximate the beta-Jacobi ensembles by the beta-Laguerre ensembles through measuring the variation distance between the eigenvalues in the two ensembles (Theorem 6 in Section 3). We then use the known results for the beta-Laguerre ensemble to get those in the beta-Jacobi ensembles.

Now we would like to mention some future problems. Notice that all the theorems above are based on the restriction (1.3). We think it could be relaxed in some situations. One possible way is that, instead of using the *uniform* approximation in (3.1), one can treat case by case for the statistics concerned in the above theorems. For example, to improve Theorem 1, one could directly evaluate the moments $\sum_{i=1}^n \lambda_i^k$ for $k \geq 1$ by computing the integration with respect to the density function $f_{\beta, a_1, a_2}(\lambda)$ in (1.1), and then check what restrictions on a_1, a_2 and n can make the integral close to the corresponding quantity in the Laguerre case.

Two other problems are discussed in Remarks 3.1 and 3.2.

Finally we give the outline of this paper. In Section 2, some known conclusions and some results on the beta-Laguerre ensembles are reviewed and proved, respectively. They will be used in the proofs for the beta-Jacobi ensembles. In Section 3, an approximation theorem for the Jacobi ensembles by the Laguerre ensemble is derived. In Section 4, we prove the main results stated in this section. In Section 5, some known and useful results are collected for our proofs.

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 $a > \frac{\beta}{2}(n-1)$. The β -Laguerre (Wishart) ensembles have the following joint probability density functions:

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

for all $\lambda_1 > 0, \dots, \lambda_n > 0$, where

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

One can see [25] for constructions of matrices to generate eigenvalues with such distributions. 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 $\lambda = (\lambda_1, \dots, \lambda_n)$ of $\mathbf{X}^* \mathbf{X}$ with $a = \frac{\beta}{2}m$ for $\beta = 1, 2$, or 4. See [30, 39, 56] for the cases $\beta = 1$ and 2, and [54] for $\beta = 4$, or (4.5) and (4.6) from [31].

It is easy to see from Theorem A.3 in Appendix that the following is true.

LEMMA 2.1 *Let $\lambda = (\lambda_1, \dots, \lambda_n)$ be random variables with the density function as in (2.1). If*

$n \rightarrow \infty$, $a \rightarrow \infty$ and $n\beta/(2a) \rightarrow \gamma \leq 1$, then

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

A large deviation result in [37, 38] implies the following ‘‘law of large numbers’’.

LEMMA 2.2 *Let $\lambda_1, \dots, \lambda_n$ be random variables with the density as in (2.1). Assume $n\beta/(2a) \rightarrow \gamma \in (0, 1]$, 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 $i = 1, 2, \dots, n$. Then μ_n converges weakly to the distribution μ_∞ with density $f_\gamma(x)$ as in (1.5) almost surely. Moreover, $\liminf_{n \rightarrow \infty} Y_{max}(n) \geq \gamma_{max}$ a.s. and $\limsup_{n \rightarrow \infty} Y_{min}(n) \leq \gamma_{min}$ a.s., where $Y_{max}(n) = \max\{Y_1, \dots, Y_n\}$ and $Y_{min}(n) = \min\{Y_1, \dots, Y_n\}$.*

The first part of the above theorem is also obtained in [24].

Proof of Lemma 2.2. From (2.1), it is obvious that the joint density function of $\{Y_i, 1 \leq i \leq n\}$ is

$$g(y_1, \dots, y_n) = Const \cdot \prod_{1 \leq i < j \leq n} |y_i - y_j|^\beta \prod_{i=1}^n y_i^{a-p} e^{-(n\beta/2\gamma) \sum_{i=1}^n y_i}$$

for $y_1 > 0, \dots, y_n > 0$. Take $\gamma(n) = a - p$, $p(n) = n$, $Q(t) = \beta t / (2\gamma)$ in Theorem A.1. Then $p(n)/n = 1$ and $\gamma(n)/n = (a - p)/n \rightarrow \beta(\gamma^{-1} - 1)/2$ as $n \rightarrow \infty$ since $p = 1 + \beta(n - 1)/2$. According to Theorem A.1, μ_n satisfies the large deviation principle with speed $\{1/n^2; n \geq 1\}$ and rate function

$$\begin{aligned} I(\nu) &= -\frac{\beta}{2} \iint \log|x - y| d\nu(x) d\nu(y) + \int \left(\frac{\beta x}{2\gamma} - \frac{\beta(\gamma^{-1} - 1)}{2} \log x \right) d\nu(x) + B \\ &= B + \frac{\beta}{\gamma^2} \left(-\frac{\gamma^2}{2} \iint \log|x - y| d\nu(x) d\nu(y) + \frac{\gamma}{2} \int (x - (1 - \gamma) \log x) d\nu(x) \right) \end{aligned}$$

where B is a finite constant. By Theorem A.2, the unique measure μ_∞ on $[0, \infty)$ to achieve the minimum of $I(\nu)$ over all probability measures on $[0, \infty)$ is the Marchenko-Pastur law with density function $f_\gamma(x)$ as in (1.5).

For $\epsilon > 0$, let $F = \{\nu; d(\nu, \mu_\infty) \geq \epsilon\}$, where d is as in (1.4). Then, F is a closed set in the weak topology. By the large deviation upper bound, there exists a constant $C > 0$ such that $P(d(\mu_n, \mu_\infty) > \epsilon) \leq e^{-n^2 C}$ as n is large enough. By the Borel-Cantelli lemma, $d(\mu_n, \mu_\infty) \rightarrow 0$ a.s. as $n \rightarrow \infty$. The first part of the conclusion is proved.

For any integer $k \geq 1$, it is easy to see

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

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

$$\begin{aligned}\liminf_{n \rightarrow \infty} Y_{max}(n) &\geq \left(\int_{\gamma_{min}}^{\gamma_{max}} y^k f_\gamma(y) dy \right)^{1/k} \quad a.s. \quad \text{and} \\ \liminf_{n \rightarrow \infty} \frac{1}{Y_{min}(n)} &\geq \left(\int_{\gamma_{min}}^{\gamma_{max}} y^{-k} f_\gamma(y) dy \right)^{1/k} \quad a.s.\end{aligned}$$

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

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

In what follows, the notation $\chi^2(s)$ stands for the χ^2 distribution with degrees of freedom s ; $\chi(s)$ denotes a positive random variables with $(\chi(s))^2$ following the $\chi^2(s)$ distribution.

LEMMA 2.3 *Let X have the Gamma distribution with density function*

$$f(x|\alpha, \theta) = \begin{cases} \frac{x^{\alpha-1} e^{-x/\theta}}{\Gamma(\alpha) \theta^\alpha}, & \text{if } x \geq 0; \\ 0, & \text{otherwise,} \end{cases}$$

where $\alpha > 0$ and $\theta > 0$ are constants. Given $b > 0$ and $\theta > 0$, set $g(\alpha) := P(X \geq b)$. Then $g(\alpha)$ is increasing over $(0, \infty)$. In particular, $P(\chi^2(s) \geq b) \leq P(\chi^2(t) \geq b)$ for any $0 < s < t$ and $b > 0$.

Proof. Since $\chi^2(s)$ has probability density function $f(x|\alpha, \theta)$ with $\alpha = p/2$ and $\theta = 2$, we only need to prove the first part of the theorem.

Obviously, X/θ has density function $f(x|\alpha, 1)$, without loss of the generality, we prove the conclusion by assuming $\theta = 1$. First, noticing $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ for $\alpha > 0$, then

$$g_1(\alpha) := g(\alpha + 1) = \frac{\int_b^\infty x^\alpha e^{-x} dx}{\int_0^\infty x^\alpha e^{-x} dx}$$

for $\alpha > -1$. Use the fact that $\frac{dx^\alpha}{d\alpha} = x^\alpha \log x$ for any $x > 0$ and $\alpha \in \mathbb{R}$ to obtain that

$$\begin{aligned}\frac{dg_1(\alpha)}{d\alpha} &= \frac{\int_b^\infty x^\alpha (\log x) e^{-x} dx \cdot \int_0^\infty x^\alpha e^{-x} dx - \int_b^\infty x^\alpha e^{-x} dx \cdot \int_0^\infty x^\alpha (\log x) e^{-x} dx}{\left(\int_0^\infty x^\alpha e^{-x} dx\right)^2} \\ &= \frac{\int_0^\infty x^\alpha e^{-x} dx \cdot \int_b^\infty x^\alpha e^{-x} dx}{\Gamma(\alpha + 1)^2} (h(b) - h(0)),\end{aligned}$$

where $h(b) := \int_b^\infty x^\alpha (\log x) e^{-x} dx / \int_b^\infty x^\alpha e^{-x} dx$ for $b \geq 0$. Fix $\alpha > -1$. To show $\frac{dg_1(\alpha)}{d\alpha} > 0$ for all $\alpha > -1$, it is enough to show that $h(b)$ is strictly increasing on $b \in [0, \infty)$. In fact,

$$\begin{aligned}&\frac{dh(b)}{db} \\ &= \frac{1}{\left(\int_b^\infty x^\alpha e^{-x} dx\right)^2} \left(-b^\alpha e^{-b} (\log b) \int_b^\infty x^\alpha e^{-x} dx + b^\alpha e^{-b} \int_b^\infty x^\alpha (\log x) e^{-x} dx \right) \\ &> 0\end{aligned}$$

since $\int_b^\infty x^\alpha (\log x) e^{-x} dx > (\log b) \int_b^\infty x^\alpha e^{-x} dx$ for all $b > 0$. \blacksquare

LEMMA 2.4 Let $\{a_n; n \geq 1\}$ and β be positive constants with $2a_n > (n-1)\beta$ for all $n \geq 1$. Set $\mathbb{T}_n = \{(n-i+1)\beta, 2a_n - (i-1)\beta; 1 \leq i \leq n\}$ for $n \geq 1$. Let $\{\chi(t), t \in \mathbb{T}_n; n \geq 1\}$ be a set of random variables defined on the same probability space. Define

$$\delta_n = \max_{t \in \mathbb{T}_n} |\chi(t) - \sqrt{t}| \quad \text{and} \quad \rho_n = \max_{t \in \mathbb{T}_n} |\chi^2(t) - t| \quad (2.3)$$

for $n \geq 1$. If $n\beta/(2a_n) \rightarrow \gamma \in (0, 1]$, then $\delta_n/\sqrt{n} \rightarrow 0$ a.s. and $\rho_n/n \rightarrow 0$ a.s. as $n \rightarrow \infty$.

Proof. First, noticing $|\mathbb{T}_n| \leq 2n$ and $\max\{t; t \in \mathbb{T}_n\} \leq n\beta + 2a_n$ for all $n \geq 1$. Since $n\beta/(2a_n) \rightarrow \gamma$ as $n \rightarrow \infty$, there exists a constant $C \geq 1$ such that $\max\{t; t \in \mathbb{T}_n\} \leq Cn$ for all $n \geq 1$. Then, for any $\epsilon > 0$,

$$\begin{aligned} P(\delta_n \geq \sqrt{n\epsilon}) &\leq 2n \cdot \max_{t \in \mathbb{T}_n} P(|\chi(t) - \sqrt{t}| \geq \sqrt{n\epsilon}) \\ &\leq 2n \cdot \max_{0 \leq t \leq Cn} P(|\chi(t) - \sqrt{t}| \geq \sqrt{n\epsilon}) \end{aligned} \quad (2.4)$$

for all $n \geq 1$. Now, taking $p_n = \lceil n^{1/3} \rceil$ for $n \geq 1$, by Lemma 2.3,

$$\max_{0 \leq t \leq p_n} P(|\chi(t) - \sqrt{t}| \geq \sqrt{n\epsilon}) \leq \max_{0 \leq t \leq p_n} P(\chi(t) \geq \sqrt{n\epsilon} - \sqrt{p_n}) \leq P\left(|\chi(p_n) - \sqrt{p_n}| \geq \frac{\sqrt{n\epsilon}}{2}\right)$$

as n is large enough. This and (2.4) imply that, given $\epsilon > 0$,

$$P(\delta_n \geq \sqrt{n\epsilon}) \leq 2n \cdot \max_{p_n \leq t \leq Cn} P\left(|\chi(t) - \sqrt{t}| \geq \frac{\sqrt{n\epsilon}}{2}\right) \quad (2.5)$$

as n is sufficiently large. Now, the last probability is equal to

$$\begin{aligned} &P\left(\chi(t) \geq \sqrt{t} + \frac{\sqrt{n\epsilon}}{2}\right) + P\left(\chi(t) \leq \sqrt{t} - \frac{\sqrt{n\epsilon}}{2}\right) \\ &\leq P\left(\chi(\lfloor t \rfloor + 1) \geq \sqrt{t} + \frac{\sqrt{n\epsilon}}{2}\right) + P\left(\chi(\lfloor t \rfloor) \leq \sqrt{t} - \frac{\sqrt{n\epsilon}}{2}\right) \quad (\text{by Lemma 2.3}) \\ &\leq P\left(\chi(\lfloor t \rfloor + 1) - \sqrt{\lfloor t \rfloor + 1} \geq -1 + \frac{\sqrt{n\epsilon}}{2}\right) + P\left(\chi(\lfloor t \rfloor) - \sqrt{\lfloor t \rfloor} \leq 1 - \frac{\sqrt{n\epsilon}}{2}\right) \\ &\leq 2P \max_{\lfloor t \rfloor \leq k \leq \lfloor t \rfloor + 1} P(|\chi(k) - \sqrt{k}| \geq \frac{\sqrt{n\epsilon}}{3}) \end{aligned}$$

as n is sufficiently large, where the inequality $\sqrt{1 + \lfloor t \rfloor} - 1 \leq \sqrt{t} \leq 1 + \sqrt{\lfloor t \rfloor}$ for all $t \geq 1$ is used in the last step. Since $|\chi(k) - \sqrt{k}| \leq |\chi^2(k) - k|/\sqrt{k}$ for any $k \geq 1$, by (2.5),

$$P(\delta_n \geq \sqrt{n\epsilon}) \leq 4n \cdot \max_{p_n \leq k \leq 2Cn} P\left(\frac{|\chi^2(k) - k|}{\sqrt{k}} \geq \frac{\sqrt{n\epsilon}}{3}\right) \quad (2.6)$$

as n is sufficiently large. Lemma 2.4 from [41] says that

$$P\left(\frac{|\sum_{i=1}^n (\xi_i^2 - 1)|}{\sqrt{n}} \geq c\right) \leq 2e^{-c^2/6}$$

for any $n \geq 1$ and $c \in (0, \sqrt{n}/2)$, where $\xi_1, \xi_2, \dots, \xi_n$ are i.i.d. random variables with $\xi_1 \sim N(0, 1)$. Since $\sqrt{n\epsilon}/3 \geq \sqrt{p_n}/3$ as n is large enough, the above inequality implies that

$$\begin{aligned} \max_{p_n \leq k \leq 2Cn} P\left(\frac{|\chi^2(k) - k|}{\sqrt{k}} \geq \frac{\sqrt{n\epsilon}}{3}\right) &\leq \max_{p_n \leq k \leq 2Cn} P\left(\frac{|\chi^2(k) - k|}{\sqrt{k}} \geq \frac{\sqrt{p_n}}{3}\right) \\ &\leq 2e^{-n^{1/3}/55} \end{aligned}$$

Recall the Geršgorin theorem: each eigenvalue of an $n \times n$ matrix $A = (a_{ij})$ lies in at least one of the disks $\{z \in \mathbb{C}; |z - a_{jj}| \leq \sum_{i \neq j} |a_{ij}|\}$ for $j = 1, 2, \dots, n$. Then

$$\lambda_{max}(n) \leq \max\{U^+, V^+, W_i^+; 2 \leq i \leq n-1\} \quad (2.9)$$

where

$$\begin{aligned} U^+ &= \chi_{2a}^2 + \chi_{2a}\chi_{\beta(n-1)}, \quad V^+ = \chi_{\beta}^2 + \chi_{2a-\beta(n-1)}^2 + \chi_{\beta}\chi_{2a-\beta(n-2)} \quad \text{and} \\ W_i^+ &= \chi_{\beta(n-i+1)}^2 + \chi_{2a-(i-1)\beta}^2 + \chi_{\beta(n-i+1)}\chi_{2a-(i-2)\beta} + \chi_{\beta(n-i)}\chi_{2a-(i-1)\beta} \end{aligned} \quad (2.10)$$

for $2 \leq i \leq n-1$, and

$$\lambda_{min}(n) \geq \min\{U^-, V^-, W_i^-; 2 \leq i \leq n-1\} \quad (2.11)$$

where

$$\begin{aligned} U^- &= \chi_{2a}^2 - \chi_{2a}\chi_{\beta(n-1)}, \quad V^- = \chi_{\beta}^2 + \chi_{2a-\beta(n-1)}^2 - \chi_{\beta}\chi_{2a-\beta(n-2)} \quad \text{and} \\ W_i^- &= \chi_{\beta(n-i+1)}^2 + \chi_{2a-(i-1)\beta}^2 - \chi_{\beta(n-i+1)}\chi_{2a-(i-2)\beta} - \chi_{\beta(n-i)}\chi_{2a-(i-1)\beta}. \end{aligned}$$

Set

$$\begin{aligned} Q_{n,i}^+ &= \beta(n-i+1) + 2a - (i-1)\beta + \sqrt{\beta(n-i+1)(2a-(i-2)\beta)} \\ &\quad + \sqrt{\beta(n-i)(2a-(i-1)\beta)} \end{aligned}$$

for $2 \leq i \leq n-1$ and $n \geq 3$. Reviewing the notation δ_n in Lemma 2.4, we have that

$$\begin{aligned} &|\chi_{\beta(n-i+1)}\chi_{2a-(i-2)\beta} - \sqrt{\beta(n-i+1)(2a-(i-2)\beta)}| \\ &= \left| \left(\chi_{\beta(n-i+1)} - \sqrt{\beta(n-i+1)} \right) \left(\chi_{2a-(i-2)\beta} - \sqrt{2a-(i-2)\beta} \right) \right. \\ &\quad \left. + \sqrt{\beta(n-i+1)} \cdot \left(\chi_{2a-(i-2)\beta} - \sqrt{2a-(i-2)\beta} \right) \right. \\ &\quad \left. + \sqrt{2a-(i-2)\beta} \left(\chi_{\beta(n-i+1)} - \sqrt{\beta(n-i+1)} \right) \right| \\ &\leq \delta_n^2 + (\sqrt{\beta n} + \sqrt{2a})\delta_n \end{aligned} \quad (2.12)$$

uniformly for all $2 \leq i \leq n-1$ and $n \geq 3$. Moreover,

$$|\chi_{\beta(n-i+1)}^2 - \beta(n-i+1)| \leq \rho_n \quad \text{and} \quad |\chi_{2a-(i-1)\beta}^2 - (2a-(i-1)\beta)| \leq \rho_n \quad (2.13)$$

for all $i \in \mathbb{T}_n$ and $n \geq 1$. Thus,

$$\left| \max_{2 \leq i \leq n-1} \{W_i^+\} - \max_{2 \leq i \leq n-1} \{Q_{n,i}^+\} \right| \leq \max_{2 \leq i \leq n-1} |W_i^+ - Q_{n,i}^+| \leq K(\rho_n + \delta_n^2 + \delta_n\sqrt{n}) \quad (2.14)$$

as n is sufficiently large, K is a constant not depending on i or n . Evidently, $Q_{n,i}^+ \leq \beta n + 2a + 2\sqrt{2a\beta n} = (\sqrt{\beta n} + \sqrt{2a})^2$ uniformly for all $i \in \mathbb{T}_n$ and $n \geq 3$. By Lemma 2.4 and the condition that $n\beta/(2a) \rightarrow \gamma$, we have

$$\limsup_{n \rightarrow \infty} \frac{\max_{2 \leq i \leq n-1} \{W_i^+\}}{n} \leq \beta(1 + \sqrt{\gamma^{-1}})^2 \quad a.s. \quad (2.15)$$

By the same but easier argument, the above is also true when $\max_{2 \leq i \leq n-1} \{W_i^+\}$ is replaced by U^+ and V^+ , respectively. Therefore, we conclude (2.7) from (2.9) and (2.15).

Now we prove (2.8). Write $(2a - (i - 1))\beta = -\beta + (2a - (i - 2))\beta$. It is easy to see

$$\begin{aligned} Q_{n,i}^- &= \beta(n - i + 1) + (2a - (i - 1))\beta - \sqrt{\beta(n - i + 1)(2a - (i - 2))\beta} \\ &\quad - \sqrt{\beta(n - i)(2a - (i - 1))\beta} \\ &\geq -\beta + \beta(n - i + 1) + (2a - (i - 2))\beta - 2\sqrt{\beta(n - i + 1)(2a - (i - 2))\beta} \\ &= -\beta + (\sqrt{\beta(n - i + 1)} - \sqrt{2a - (i - 2)})^2 \end{aligned}$$

for $2 \leq i \leq n - 1$ and $n \geq 3$. Use the equality $\sqrt{x} - \sqrt{y} = (x - y)/(\sqrt{x} + \sqrt{y})$ to obtain

$$(\sqrt{\beta(n - i + 1)} - \sqrt{2a - (i - 2)})^2 = \frac{(\beta(n - 1) - 2a)^2}{(\sqrt{\beta(n - i + 1)} + \sqrt{2a - (i - 2)})^2},$$

which is increasing in $i \in \{2, 3, \dots, n - 1\}$. Since $a > (n - 1)/2$, we get

$$\min_{2 \leq i \leq n-1} (\sqrt{\beta(n - i + 1)} - \sqrt{2a - (i - 2)})^2 = \frac{(\beta(n - 1) - 2a)^2}{(\sqrt{\beta(n - 1)} + \sqrt{2a})^2}.$$

for $n \geq 3$. Combining all the above steps, and using the condition that $n\beta/(2a) \rightarrow \gamma$, we arrive at

$$\liminf_{n \rightarrow \infty} \frac{\min_{2 \leq i \leq n-1} \{Q_{n,i}^-\}}{n} \geq \frac{\beta(1 - \gamma^{-1})^2}{(1 + \sqrt{\gamma^{-1}})^2} = \beta(1 - \sqrt{\gamma^{-1}})^2 \quad a.s. \quad (2.16)$$

By (2.12) and (2.13) and the same argument as in (2.14), we obtain

$$\left| \min_{2 \leq i \leq n-1} \{W_i^-\} - \min_{2 \leq i \leq n-1} \{Q_{n,i}^-\} \right| \leq \max_{2 \leq i \leq n-1} |W_i^- - Q_{n,i}^-| \leq K(\rho_n + \delta_n^2 + \sqrt{n}\delta_n)$$

as n is sufficiently large. This together with (2.16) and Lemma 2.4 concludes

$$\liminf_{n \rightarrow \infty} \frac{\min_{2 \leq i \leq n-1} \{W_i^-\}}{n} \geq \beta(1 - \sqrt{\gamma^{-1}})^2 \quad a.s.$$

Similarly,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{U^-}{n} &= \lim_{n \rightarrow \infty} \frac{2a - \sqrt{2a\beta(n - 1)}}{n} = \beta\gamma^{-1} - \beta\sqrt{\gamma^{-1}} \quad a.s. \\ \lim_{n \rightarrow \infty} \frac{V^-}{n} &= \lim_{n \rightarrow \infty} \frac{2a - \beta(n - 1)}{n} = \beta\gamma^{-1} - \beta \quad a.s. \end{aligned}$$

Since $\gamma \in (0, 1]$, it is obvious that $\min\{\beta\gamma^{-1} - \beta\sqrt{\gamma^{-1}}, \beta\gamma^{-1} - \beta\} \geq \beta(1 - \sqrt{\gamma^{-1}})^2$. By (2.11), the above three assertions imply (2.8). \blacksquare

3 An Approximation Result

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

$$\|\mu - \nu\| = 2 \cdot \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)| = \int_{\mathbb{R}^m} |f(x) - g(x)| dx_1 \cdots dx_m \quad (3.1)$$

if μ and ν have density functions $f(x)$ and $g(x)$ with respect to the Lebesgue measure. For a random variable Z , we use $\mathcal{L}(Z)$ to denote its probability distribution. The following is the key tool to prove the results stated in Introduction.

THEOREM 6 *Let $\mu = (\mu_1, \dots, \mu_n)$ and $\lambda = (\lambda_1, \dots, \lambda_n)$ 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 $\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| \rightarrow 0$.*

REMARK 3.1 *The condition (1.3) is actually sharp for $\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| \rightarrow 0$ for $\beta = 1$ and 2. In fact, let U be an $N \times N$ random unitary matrix with real ($\beta = 1$) and complex ($\beta = 2$), chosen with Haar measure. Decompose*

$$U = \begin{pmatrix} A_{n_1, n_2} & C_{n_1 \times (N-n_2)} \\ B_{(N-n_1) \times n_2} & D_{(N-n_1) \times (N-n_2)} \end{pmatrix}.$$

Assume $n_1 \geq n_2$ and $n_1 + n_2 \leq N$. Set $Y = A_{n_1, n_2}^* A_{n_1, n_2}$, we see from (3.15) in [34] that the eigenvalues $\lambda_1, \dots, \lambda_{n_2}$ of Y has density

$$\text{Const} \cdot \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta \cdot \prod_{i=1}^n \lambda_i^{\alpha_1 - p} (1 - \lambda_i)^{\alpha_2 - p},$$

which belongs to the beta-Jacobi ensemble (1.1) with

$$a_1 = \frac{\beta}{2} n_1, \quad a_2 = \frac{\beta}{2} (N - n_1), \quad n = n_2 \quad \text{and} \quad p = 1 + \frac{\beta}{2} (n - 1).$$

It is shown in [43] for $\beta = 1$ and [42] for $\beta = 2$ that the variation distance d' between $\sqrt{N}A_{n_1, n_2}$ and X_n goes to zero if $n_1 = o(\sqrt{N})$ and $n_2 = o(\sqrt{N})$, where X_n is the joint distribution of $n_1 n_2$ independent and identically distributed real Gaussian ($\beta = 1$) or complex Gaussian ($\beta = 2$) random variables. The orders $n_1 = o(\sqrt{N})$ and $n_2 = o(\sqrt{N})$, which correspond to that $a_1 = o(\sqrt{a_2})$ and $n_1 = o(\sqrt{a_2})$ in (1.3), are also proved to be sharp for both cases in [42, 43]. A further analysis shows that $d' \rightarrow 0$ if and only if $\|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| \rightarrow 0$. This tells us that the orders in Theorem 6 are sharp for $\beta = 1$ and 2. However, it is not known whether the same remains true for other $\beta > 0$.

REMARK 3.2 *The condition “ $n\beta/(2a_1) \rightarrow \gamma$ ” in (1.3) is required in Theorem 6. In the same contexts as in [43] for $\beta = 1$ and [42] for $\beta = 2$, the condition is not imposed. Although the condition is harmless in proving the main results stated in Introduction, it would be interesting to see if it can be removed for other $\beta > 0$.*

Now we start to prove the above theorem by developing several lemmas.

LEMMA 3.1 *Let $n \geq 2$. Let $\mu = (\mu_1, \dots, \mu_n)$ and $\lambda = (\lambda_1, \dots, \lambda_n)$ 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

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

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

Proof. It is enough to show

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

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

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

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

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

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

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

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

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

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,

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

LEMMA 3.2 Let $h(x) = x \log x$ for $x > 0$. For a fixed constant $\beta > 0$, an integer $n \geq 1$ 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 $n \rightarrow \infty$, $a_1 \rightarrow \infty$ and $a_2 \rightarrow \infty$ in a way that $a_1 = o(\sqrt{a_2})$ and $n = o(\sqrt{a_2})$, then

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

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,

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

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

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

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

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

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

$$n \rightarrow \infty, a_1 \rightarrow \infty \text{ and } a_2 \rightarrow \infty \text{ such that } a_1 = o(\sqrt{a_2}), \text{ and } n = o(\sqrt{a_2}). \quad (3.8)$$

Now,

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

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

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

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

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

The conclusion follows because the last two terms are all of order $o(1)$ as (3.8) holds. ■

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

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

Proof. We claim that it suffices to prove

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

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

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

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

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

as $x = \operatorname{Re}(z) \rightarrow +\infty$, where $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ with $\operatorname{Re}(z) > 0$, see, e.g., p.368 from [35] or (37) on p.204 from [3]. It follows that

$$\log K_n \quad (3.11)$$

$$= -\alpha n b_1 \log(\alpha b_2) + \sum_{i=1}^n \{ \alpha(b_1 + b_2 - i + 1) \log \alpha(b_1 + b_2 - i + 1) - \alpha(b_2 - i + 1) \log \alpha(b_2 - i + 1) \} \quad (3.12)$$

$$-\alpha n b_1 - \frac{1}{2} \sum_{i=1}^n \log \frac{b_1 + b_2 - i + 1}{b_2 - i + 1} + O\left(\frac{1}{b_2 - n}\right) \quad (3.13)$$

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

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

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

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

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

$$\log K_n = \alpha nb_1 \cdot \frac{b_1 - n}{2b_2} + o(1) = \frac{na_1(a_1 - \frac{\beta}{2}n)}{2a_2} + o(1)$$

as (3.8) holds. This gives (3.10). \blacksquare

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

LEMMA 3.4 *Let $\mu = (\mu_1, \dots, \mu_n)$ 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*

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

in probability as $n \rightarrow \infty$.

Proof. From (3.3), we see that

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

By Lemma 2.5, since $\frac{n\beta}{2a_1} \rightarrow \gamma \in (0, 1]$ by (1.3),

$$\frac{\max_{1 \leq i \leq n} \mu_i}{n} \rightarrow \beta(1 + \sqrt{\gamma^{-1}})^2 \quad (3.15)$$

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

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

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

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

This is because, for any two sequences random variables $\{\xi_n; n \geq 1\}$ and $\{\eta_n; n \geq 1\}$, if $\xi_n \rightarrow 1$ in probability and $P(\xi_n \neq \eta_n) \rightarrow 0$ as $n \rightarrow \infty$, then $\eta_n \rightarrow 1$ in probability as $n \rightarrow \infty$. 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 $\Omega_n := \{\max_{1 \leq i \leq n} \mu_i \leq \delta_n\}$. Noticing $\log(1-x) = -x - (x^2/2) + O(x^3)$ as $x \rightarrow 0$,

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

on Ω_n . Now, on Ω_n again,

$$\frac{1}{a_2^3} \sum_{i=1}^n \mu_i^3 \leq \frac{n(\delta_n)^3}{a_2^3} = \left(\frac{\delta_n}{\sqrt{a_2}}\right)^3 \cdot \frac{n}{\sqrt{a_2}} \cdot \frac{1}{a_2} \rightarrow 0 \quad (3.19)$$

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 ,

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

as (1.3) holds. By Lemma 2.1, as $n \rightarrow \infty$,

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

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

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

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

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

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

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

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

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

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

as (1.3) holds. Thus, on Ω_n ,

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

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

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

Proof. By the Skorohod representation theorem (see, e.g., p.85 from [26]), w.l.o.g., assume $X_n \rightarrow 1$ almost surely as $n \rightarrow \infty$. Thus, for any $K > 1$, we have $X_n I(X_n \leq K) \rightarrow 1$ almost surely as $n \rightarrow \infty$. This gives that $E|X_n I(X_n \leq K) - 1| \rightarrow 0$ and $EX_n I(X_n \leq K) \rightarrow 1$ as $n \rightarrow \infty$. The second assertion and the condition that $EX_n \rightarrow 1$ imply $EX_n I(X_n > K) \rightarrow 0$ as $n \rightarrow \infty$. Therefore,

$$E|X_n - 1| \leq E|X_n I(X_n \leq K) - 1| + EX_n I(X_n > K) \rightarrow 0$$

as $n \rightarrow \infty$. \blacksquare

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

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

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,

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

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

4 The Proofs of Main Results

By using Theorem 6 developed in Section 3, we now are ready to prove the results stated in Introduction.

Proof of Theorem 1. Set

$$\nu_n = \frac{1}{n} \sum_{i=1}^n I_{\frac{a_2}{n} \lambda'_i}$$

for $(\lambda'_1, \dots, \lambda'_n) \in [0, +\infty)^n$. Then, recall the definition of d in (1.4), by the triangle inequality,

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

where the Lipschitz inequality $|f(x) - f(y)| \leq 1$ is used in the last step. This says that $d(\mu_n, \mu_0)$, as a function of $(\lambda_1, \dots, \lambda_n)$, is continuous for each $n \geq 2$, and is therefore Borel-measurable. From (3.1) we see that

$$P(d(\mu_n, \mu_0) \geq \epsilon) \leq P(d(\mu'_n, \mu_0) \geq \epsilon) + \|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\| \quad (4.1)$$

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

$$\frac{1}{n} \sum_{i=1}^n I_{\frac{\mu_i \gamma}{n\beta}} \text{ converges weakly to } \mu_\infty \quad (4.2)$$

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, $d(\mu'_n, \mu_0) \rightarrow 0$ almost surely. This, (4.1) and Theorem 6 in Section 3 prove the theorem. ■

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

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

for any $\epsilon > 0$, where $\mu = (\mu_1, \dots, \mu_n)$ has density $f_{\beta, a_1}(\mu)$ as in (2.1) with $a = a_1$ and $n\beta/2a_1 \rightarrow \gamma \in (0, 1]$. From Lemma 2.5, we know $\mu_{max}(n)/(2n) \rightarrow \beta(1 + \sqrt{\gamma^{-1}})^2/2 = \beta(1 + \sqrt{\gamma})^2/(2\gamma)$ in probability. This together with (4.3) and Theorem 6 in Section 3 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,

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

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 $i \geq 1$ (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 $n\beta/2a_1 \rightarrow \gamma \in (0, 1]$. The conclusion then follows from this, Lemma A.3 and Theorem 6 in Section 3. \blacksquare

Proof of Theorem 4. The assumption that $2\beta^{-1}a_1 - n = c$ and $n = o(\sqrt{a_2})$ imply that $n\beta/2a_1 \rightarrow 1$ and $a_1 = o(\sqrt{a_2})$. Thus the conclusion of Theorem 6 in Section 3 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 [59] correspond to “ θ_i ” and “ c ” here, respectively. By Theorem 1 from [59], for fixed integer $k \geq 1$,

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

as $n \rightarrow \infty$. By Theorem 6 in Section 3,

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

for any Borel set $B_n \subset \mathbb{R}^n$ for $n \geq 1$. Combining the above two limits we obtain

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

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

Proof of Theorem 5. Recalling (2.1), let

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

Let $(\theta_1, \dots, \theta_n)$ have density $f_{\beta, a}$ as in (2.1). Noticing, “ $\beta\lambda_i$ ” in Theorem 1.4 from [58] corresponds to “ θ_i ” here; “ κ ” in Theorem 1.4 from [58] is equal to $2\beta^{-1}a$, and $\frac{\beta}{2}(n-1) + 1 = p$, and the k -th lowest eigenvalue of \mathcal{H}_β is the $(n-k+1)$ -th largest eigenvalue of $-\mathcal{H}_\beta$. Then by Theorem 1.4 from [58],

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

as $n \rightarrow \infty$ and $a \rightarrow \infty$ such that n/a converges to a nonzero, finite constant. In other words,

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

for any Borel set $A \subset \mathbb{R}^k$. By Theorem 6 in Section 3, assuming (1.3) and $a = a_1$,

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

for any Borel set $B_n \subset \mathbb{R}^n$ for $n \geq 1$. Evidently, for $x = (x_1, \dots, x_m) \in \mathbb{R}^m$, the function $g(x) := x^{(l)}$, the l -th largest one in $\{x_1, \dots, x_m\}$, is a continuous function in x . Replacing a by a_1 in (4.4), the above two assertions conclude that

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

5 Appendix

Let $Q(x)$ be a real continuous function on $[0, \infty)$ such that for any $\epsilon > 0$

$$\lim_{x \rightarrow +\infty} x e^{-\epsilon Q(x)} = 0. \quad (5.1)$$

For each integer $n \geq 1$, let $p(n)$ be an integer depending on n . Assume the joint probability density of $(\lambda_1, \dots, \lambda_{p(n)})$ is of the form

$$\nu_n := \frac{1}{Z_n} \exp \left(-n \sum_{i=1}^{p(n)} Q(t_i) \right) \prod_{i=1}^{p(n)} t_i^{\gamma(n)} \prod_{1 \leq i < j \leq p(n)} |t_i - t_j|^\beta \quad (5.2)$$

where $\beta > 0$ is fixed and $\gamma(n) \geq 0$ depends on n . Let μ_n be the empirical probability measure of $\lambda_1, \dots, \lambda_{p(n)}$. Under the weak topology, the large deviations for $\{\mu_n\}$ is given below. For a reference of general large deviations, one can see, e.g., [15] and [17].

THEOREM A.1 (Theorem 1 in [38]) *Assume $p(n)/n \rightarrow \kappa \in (0, 1]$ and $\gamma(n)/n \rightarrow \tau \geq 0$ as $n \rightarrow \infty$. Then the finite limit $B := \lim_{n \rightarrow \infty} n^{-2} \log Z_n$ exists and $\{\mu_n; n \geq 1\}$ satisfies the large deviation principle with speed $\{n^{-2}; n \geq 1\}$ and good rate function*

$$I(\nu) := -\frac{\kappa^2 \beta}{2} \iint \log |x - y| d\nu(x) d\nu(y) + \kappa \int (Q(x) - \tau \log x) d\nu(x) + B \quad (5.3)$$

for all probability measure ν defined on $[0, \infty)$. Moreover, there exists a unique probability measure ν_0 on $[0, \infty)$ such that $I(\nu_0) = 0$.

In Theorem 1 from [38] or Theorem 5.5.1 from [37], the limit τ above is required to be strictly positive. However, after a check, it is found that the conclusion also holds for $\tau = 0$.

THEOREM A.2 (part of Theorem 8 from [38]) *Let $\gamma \in (0, 1]$, $\gamma_{min} = (1 - \sqrt{\gamma})^2$ and $\gamma_{max} = (1 + \sqrt{\gamma})^2$. For probability measure ν , define*

$$J(\nu) = -\frac{\gamma^2}{2} \iint \log |x - y| d\nu(x) d\nu(y) + \frac{\gamma}{2} \int (x - (1 - \gamma) \log x) d\nu(x).$$

Then the unique minimizer of $J(\nu)$ over all probability measures on $[0, +\infty)$ is the Marchenko-Pastur law with density function $f_\gamma(x)$ as in (1.5).

Proof. Take a in [38] equal to γ . Notice

$$\sqrt{4\gamma - (x - 1 - \gamma)^2} = \sqrt{(x - (1 - \sqrt{\gamma}^2))((1 + \sqrt{\gamma})^2 - x)} = \sqrt{(x - \gamma_{min})(\gamma_{max} - x)}$$

for all $x \in [\gamma_{min}, \gamma_{max}]$. Also, over all probability measure ν on $[0, \infty)$, taking the minimum for

$$\begin{aligned} I(\nu) : &= -\frac{\gamma^2}{2} \int \int \log|x - y| d\nu(x) d\nu(y) + \frac{\gamma}{2} \int (x - (1 - \gamma) \log x) d\nu(x) \\ &\quad - \frac{1}{4} (3\gamma - \gamma^2 \log \gamma + (1 - \gamma)^2 \log(1 - \gamma)) \end{aligned}$$

is the same as doing so for $J(\nu)$. Then the conclusion follows from Theorem 8 from [38]. \blacksquare

THEOREM A.3 (Theorem 1.5 from [23]) Let $\lambda_1, \dots, \lambda_n$ be random variables with the density function as in (2.1). Assume $n\beta/(2a) \rightarrow \gamma \leq 1$. Define

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

for $i \geq 1$. Then, as $n \rightarrow \infty$, (X_1, \dots, X_k) converges weakly to a normal distribution $N_k(\mu, \Sigma)$ for some μ and Σ .

References

- [1] Absil, P. A., Edelman, A. and Koev, P. (2006). On the largest principal angle between random subspaces. *Linear Algebra and its Applications* 414, 288-294.
- [2] Adler, M. and Van Moerbeke, P. (2001). Integrals over classical groups, random permutations, Toda and Toeplitz lattices. *Comm. Pure Appl. Math.* 54, 153-205.
- [3] Ahlfors, L. V. (1979). *Complex Analysis*, 3rd. Edition. McGraw-Hill, Inc.
- [4] Aomoto, K. (1987). Jacobi polynomials associated with Selberg integrals. *SIAM (Soc. Ind. Appl. Math.) J. Math. Anal.* 18, 545-549.
- [5] Anderson, G., Guionnet, A. and Zeitouni, O. *An Introduction to Random Matrices* (Cambridge Studies in Advanced Mathematics). To appear.
- [6] Anderson, T. (1984). *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, Second edition.
- [7] Bai, Z. D. and Silverstein, J. W. (2006). *Spectral Analysis of Large Dimensional Random Matrices*. Science Press.
- [8] Baik, J., Deift, P. and Johansson, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *Jour. Amer. Math. Soc.* 12, 1119-1178.
- [9] Baker, T. and Forrester, P. (1997). The Calogero-Sutherland model and generalized classical polynomials. *Commun. Math. Phys.* 188, 175-216.
- [10] Byrc, W., Far, R., Oraby, T. and Speicher, R. (2008). On slow-fading MIMO systems with non-separable correlation. *IEEE Transactions on Information Theory* 54, 544-553.
- [11] Bryc, W., Dembo, A. and Jiang, T. (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.* 34, 1-38.

- [12] Collins, B. (2005). Product of random projections, Jacobi ensembles and universality problems arising from free probability. *Probability Theory and Related Fields*, 133, 315-344.
- [13] Collins, B. (2003). *Intégrales Matricielles et Probabilités Non-commutatives*. Thèse de Doctorat of Université Paris 6.
- [14] Constantine, A. (1963). Some non-central distribution problems in multivariate analysis. *Ann. Math. Stat.* 34, 1270-1285.
- [15] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. Springer, 2nd edition.
- [16] Demni, N. β -Jacobi processes. http://arxiv.org/PS_cache/arxiv/pdf/0901/0901.0324v3.pdf.
- [17] Deuschel, J. and Stroock, D. (1989). *Large Deviations*. Academic Press.
- [18] Diaconis, P. and Evans, S. (2001). Linear functionals of eigenvalues of random matrices. *Transactions Amer. Math. Soc.* 353, 2615-2633.
- [19] Diaconis, P., Eaton, M. and Lauritzen, L. (1992). Finite deFinetti theorem in linear models and multivariate analysis. *Scand. J. Statist.* 19(4), 289-315.
- [20] Diaconis, P. and Shahshahani, M. (1994). On the Eigenvalues of Random Matrices. In *Studies in Applied Probability*; Gani, J., Ed.; Jour. Appl. Probab.: Special Vol. 31A, 49-62.
- [21] Dudley, R. (2002). *Real Analysis and Probability*. Cambridge University Press.
- [22] Dumitriu, I. and Koev, P. (2003). Distributions of the extreme eigenvalues of Beta-Jacobi random matrices. *SIAM Journal on Matrix Analysis and Applications*, 30(1), 1-6.
- [23] Dumitriu, I. and Edelman, A. (2006). Global spectrum fluctuations for the β -Hermite and β -Laguerre ensembles via matrix models. *Journal of Mathematical Physics* 47, 063302.
- [24] Dumitriu, I. (2003). *Eigenvalue Statistics for Beta-Ensembles*. PhD thesis, Massachusetts Institute of Technology.
- [25] Dumitriu, I. and Edelman, A. (2002). Matrix models for beta-ensembles. *J. Math. Phys.*, 43(11), 5830-5847.
- [26] Durrett, R. (1996). *Probability: Theory and Examples*. Duxbury Press, 2nd Ed.
- [27] Dyson, F. (1962). Statistical theory of the energy levels of complex systems. I. *J. Mathematical Phys.* 3, 140-156.
- [28] Eaton, M. (1983). *Multivariate Statistics: A Vector Space Approach* (Wiley Series in Probability and Statistics), John Wiley & Sons Inc.
- [29] Edelman, A. and Sutton, B. (1989). The Beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems. *Foundations of Computational Mathematics*, 8(2), 259-285.
- [30] Edelman, A. (1989). *Eigenvalues and condition numbers of random matrices*, PhD thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA.
- [31] Edelman, A. and Rao, N. (2005). Random matrix theory. *Acta Numerica*, 1-65.
- [32] Forrester, P. *Log-Gases and Random Matrices*. To appear.
- [33] Forrester, P. and Warnaar, S. (2008). The importance of the Selberg integral. *Bull. Amer. Math. Soc.* 45(4), 489-534.
- [34] Forrester, P. (2004). Quantum conductance problems and the Jacobi ensemble. *J. Phys. A: Math. Gen.* 39, 6861-6870.

- [35] Gamelin, T. W. (2001). *Complex Analysis*. 1st Ed., Springer.
- [36] Guionnet, A. (2004). Large deviations and stochastic calculus for large random matrices. *Probab. Surveys* 1, 72-172.
- [37] Hiai, F. and Petz, D. (2006). *The Semicircle Law, Free Random Variables and Entropy* (Mathematical Surveys and Monographs). American Mathematical Society.
- [38] Hiai, F. and Petz, D. (1998). Eigenvalue density of the Wishart matrix and large deviations. *Infinite Dim. Anal. and Quantum Probability* 1, 633-646.
- [39] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Stat.* 35, 475-501.
- [40] Jiang, T. (2009). The entries of Haar-invariant matrices from the classical compact groups. *Journal of Theoretical Probability* (Online with DOI 10.1007/s10959-009-0241-7).
- [41] Jiang, T. (2009). The entries of circular orthogonal ensembles. *Journal of Mathematical Physics* 50(6), 063302.
- [42] Jiang, T. (2009). Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles. *Probability Theory and Related Fields* 144(1), 221-246.
- [43] Jiang, T. (2006). How many entries of a typical orthogonal matrix can be approximated by independent normals? *Ann. Probab.* 34(4), 1497-1529.
- [44] Jiang, T. (2005). Maxima of Entries of Haar Distributed Matrices. *Probability Theory and Related Fields* 131, 121-144.
- [45] Jiang, T. (2004). The limiting distributions of eigenvalues of sample correlation matrices. *Sankhya* 66(1), 35-48.
- [46] Johansson, K. (2000). Shape fluctuation and random matrices. *Comm. Math. Phys.* 209, 437-476.
- [47] Johansson, K. (1997). On random matrices from the compact classical groups. *Ann. Math.* 145, 519-545.
- [48] Jonsson, D. (1982). Some limit theorem for eigenvalues of a sample covariance matrices. *J. Multivariate Anal.* 12, 1-38.
- [49] Johnstone, I. (2008). Multivariate analysis and Jacobi ensembles: largest eigenvalue, Tracy-Widom limits and rates of convergence. *Ann. Stat.* 36(6), 2638-2716.
- [50] Kadell, K. (1997). The Selberg-Jack polynomials. *Adv. Math.* 130, 33-102.
- [51] Killip, R. and Nenciu, I. (2004). Matrix models for circular ensembles. *Int. Math. Res. Not.* 50, 2665C2701.
- [52] Lippert, R. (2003). A matrix model for the beta-Jacobi ensemble. *J. Math. Phys.* 44(10), 4807.
- [53] Lytova, A. and Pastur, L. (2009). Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.* 37(5), 1778-1840.
- [54] Macdonald, I. (1998). *Symmetric Functions and Hall Polynomials*, Oxford Mathematical Monographs, 2nd Edn, Oxford University Press.
- [55] Marchenko, V. A. and Pastur, L. A. (1967). Distribution of some sets of random matrices. *Math. USSR-sb.* 1, 457-483.
- [56] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.
- [57] Okounkov, A. and Olshanski, G. (1997). Shifted Jack polynomials, binomial formula, and applications. *Mathematical Research Letters* 4, 69-78.

- [58] Ramírez, J., Rider, B. and Virág, B.: Beta ensembles, stochastic Airy spectrum and a diffusion. Available at http://arxiv.org/PS_cache/math/pdf/0607/0607331v3.pdf.
- [59] Ramírez, J. and Rider, B. (2009). Diffusion at the random matrix hard edge. *Comm. Math. Phys.* 288(3), 887-906.
- [60] Silverstein, J. (1985). The smallest eigenvalue of a large dimensional Wishart matrix. *Ann. Probab.* 13(4), 1364-1368.
- [61] Stanley, R. (1989). Some combinatorial properties of Jack symmetric functions. *Adv. Math.* 77, 76-115.
- [62] Tao, T. and Vu, V. Random matrices: universality of local eigenvalue statistics. To appear in *Acta Mathematica*.
- [63] Tao, T., Vu, V. and Krishnapur, M. Random Matrices: universality of ESDS and the circular law. To appear in *Ann. Probab.* Available at http://arxiv.org/PS_cache/arxiv/pdf/0807/0807.4898v5.pdf.
- [64] Tracy, C. and Widom, H. (2000). The distribution of largest eigenvalue in the Gaussian ensembles, in *Calogero-Moser-Sutherland Models*, Vol. 4 of *CRM Series in Mathematical Physics* (J.van Diejen and L. Vinet, eds), Springer, Berlin, 461-472.
- [65] Tracy, C. and Widom, H. (1996). On the orthogonal and symplectic matrix ensembles. *Comm. Math. Physics*, 177, 727-754.
- [66] Tracy, C. and Widom, H. (1994). Level-spacing distributions and Airy kernel. *Comm. Math. Physics*, 159, 151-174.
- [67] Wang, C. and Uhlenbeck, G. (1945). On the theory of the Brownian motion. II. *Rev. Modern Phys.* 17, 323-342.
- [68] Wigner, E. P. (1958). On the distribution of roots of certain symmetric matrices. *Ann. Math.* 67, 325-328.