Approximation of Beta-Jacobi Ensembles by Beta-Laguerre Ensembles

Yutao MA¹, Xinmei SHEN²

1 School of Mathematical Sciences & Laboratory of Mathematics and Complex Systems of Ministry of Education, Beijing Normal University, Beijing 100875, China
2 School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

© Peking University 2023

Abstract Consider beta-Laguerre ensembles \( \mu \) with parameters \( m, a_1 \) and beta-Jacobi ensembles \( \lambda \) with parameters \( m, a_1, a_2 \). With the help of tridiagonal models of beta ensembles, we are able to prove that \( \lim_{a_2 \to \infty} d(\mathcal{L}(2a\lambda), \mathcal{L}(\mu)) = 0 \) if \( a_1 m = o(a_2) \) and \( \lim_{a_2 \to \infty} d(\mathcal{L}(2a\lambda), \mathcal{L}(\mu)) > 0 \) if \( \lim_{a_2 \to \infty} \frac{a_1 m}{a_2} = \sigma > 0 \), by contrast, where \( a := a_1 + a_2 \) and \( d \) is total variation distance or Kullback–Leibler divergence. This result improves the approximation in [9].

Keywords Beta-Laguerre ensembles, beta-Jacobi ensembles, total variation distance, Kullback–Leibler divergence

MSC2020 60B20, 60F99

1 Introduction

Let \( \mu \) and \( \nu \) be two probability measures on \((\mathbb{R}^n, \mathcal{B})\), where \( \mathbb{R}^n \) is the \( n \)-dimensional Euclidean space and \( \mathcal{B} \) is the Borel \( \sigma \)-algebra on \( \mathbb{R}^n \). We will consider the total variation distance and Kullback–Leibler divergence between \( \mu \) and \( \nu \):

(1) Total variation distance between \( \mu \) and \( \nu \), denoted by \( \|\mu - \nu\|_{TV} \), is defined by

\[
\|\mu - \nu\|_{TV} = 2 \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)| = \int_{\mathbb{R}^n} |f(x) - g(x)| \, dx
\]

provided \( \mu \) and \( \nu \) have density functions \( f \) and \( g \) with respect to the Lebesgue measure, respectively.

Received December 9, 2020; accepted July 13, 2021
Corresponding author: Yutao MA, E-mail: mayt@bnu.edu.cn
(2) Kullback–Leibler divergence between $\mu$ and $\nu$ is defined by

$$D_{KL}(\mu||\nu) = \int_{\mathbb{R}^n} \frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu} = \int_{\mathbb{R}^n} f(x) \log \frac{f(x)}{g(x)} dx.$$ 

Let $\beta > 0$ be a constant and $m \geq 1$ be an integer. A beta-Jacobi ensemble, also called the beta-MANOVA ensemble, is a set of random variables $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m) \in [0, 1]^m$ with joint probability density function

$$f_{\beta,a_1,a_2}(x_1, \ldots, x_m) = C_{J}^{\beta,a_1,a_2} \prod_{1 \leq i < j \leq m} |x_i - x_j|^\beta \prod_{i=1}^{m} x_i^{a_1-r}(1-x_i)^{a_2-r}, \quad (1.1)$$

where $a_1, a_2 > \frac{\beta}{2}(m-1)$ and $r := 1 + \frac{\beta}{2}(m-1)$, and

$$C_{J}^{\beta,a_1,a_2} = \prod_{j=1}^{m} \frac{\Gamma(1 + \beta/2)\Gamma(a_1 + a_2 - \beta(m-j)/2)}{\Gamma(1 + \beta j/2)\Gamma(a_1 - \beta(m-j)/2)\Gamma(a_2 - \beta(m-j)/2)}.$$

The density has close connections to the multivariate analysis of variance (MANOVA). For $\beta = 1, 2, 4$, the density function $f_{\beta,a_1,a_2}$ in (1.1) is the joint probability density function of the eigenvalues of matrices $Y'Y(Y'Y + Z'Z)^{-1}$ with $a_1 = \beta n_1/2$ and $a_2 = \beta n_2/2$. Here $Y = Y_{n_1 \times m}$ and $Z = Z_{n_2 \times m}$ are independent matrices with $n_1, n_2 \geq m$ and the entries of both matrices independent random variables with the standard real, complex or quaternion Gaussian distributions. See [1] for $\beta = 1$ and [12] for $\beta = 2$, respectively.

A beta-Laguerre ensemble is a set of non-negative random variables $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ with joint density function

$$f_{\beta,a_1}(x_1, \ldots, x_m) = C_{L}^{\beta,a_1} \prod_{1 \leq i < j \leq m} |x_i - x_j|^\beta \prod_{i=1}^{m} x_i^{a_1-r}e^{-\frac{x_i}{2}}, \quad (1.2)$$

where $a_1 > \frac{\beta}{2}(m-1)$ and $r = 1 + \frac{\beta}{2}(m-1)$, and

$$C_{L}^{\beta,a_1} = 2^{-ma_1} \prod_{j=1}^{m} \frac{\Gamma(\frac{\beta}{2})}{\Gamma(1 + \frac{\beta}{2})\Gamma(a_1 - \frac{\beta(m-j)}{2})}.$$ 

It is clear that

$$\frac{f_{\beta,a_1,a_2}(x_1, \ldots, x_m)}{f_{\beta,a_1}(x_1, \ldots, x_m)} = \frac{C_{J}^{\beta,a_1,a_2}}{C_{L}^{\beta,a_1}} \prod_{i=1}^{m} (1-x_i)^{a_2-r}e^{\frac{x_i}{2}}.$$ 

Let $\Gamma_n = (\gamma_{ij})$ be a random orthogonal matrix which is uniformly distributed on the orthogonal group $O(n)$. Let $Z_n$ be the $p_n \times q_n$ upper left block of $\Gamma_n$, where $p_n$ and $q_n$ are two positive integers. Denoted by $L(\sqrt{n}Z_n)$ the joint probability distribution of the $p_nq_n$ random entries of $\sqrt{n}Z_n$ and $L(G_n)$
the joint distribution of $p_nq_n$ independent standard normals. Let $f_n$ and $g_n$ be the probability density function of $\mathcal{L}(\sqrt{n}Z_n)$ and $\mathcal{L}(G_n)$ with respect to the Lebesgue measure, respectively. According to the explicit expression of $f_n$ in [11], it has a particular form of $f;a_1,a_2$ with $\beta = \frac{1}{2}$, $m = q$, $a_1 = \frac{p}{2}$ and $a_2 = \frac{n-p}{2}$. In [8], Jiang proved that when $p_n = o(\sqrt{n})$ and $q_n = o(\sqrt{n})$, 

$$\lim_{n \to \infty} \|\mathcal{L}(\sqrt{n}Z_n) - \mathcal{L}(G_n)\|_{TV} = 0$$

and when $p_n = O(\sqrt{n})$ and $q_n = O(\sqrt{n})$, 

$$\lim_{n \to \infty} \|\mathcal{L}(\sqrt{n}Z_n) - \mathcal{L}(G_n)\|_{TV} > 0.$$ 

This is the first result to characterize exactly how many entries of a typical orthogonal matrix could be approximated by independent normals. Recently, Jiang and the first author in [11] completely resolved this problem. Precisely, they showed that 

$$\lim_{n \to \infty} d(\mathcal{L}(\sqrt{n}Z_n), \mathcal{L}(G_n)) = 0, \text{ if } pq = o(n);$$

$$\lim_{n \to \infty} d(\mathcal{L}(\sqrt{n}Z_n), \mathcal{L}(G_n)) > 0, \text{ if } pq = O(n).$$ 

Here $d$ is the total variation distance, the Kullback–Leibler divergence or the Hellinger distance. In 2013, Jiang in [9] worked on general $\beta > 0$. He proved that when 

$$m \to \infty, a_1 \to \infty \text{ and } a_2 \to \infty \text{ such that } a_1 = o(\sqrt{a_2}), m = o(\sqrt{a_2}) \text{ and } \frac{m\beta}{2a_1} \to \gamma \in (0,1],$$

it holds 

$$\lim_{a_2 \to \infty} \|\mathcal{L}(2a_2\lambda) - \mathcal{L}(\mu)\|_{TV} = 0,$$

where $\lambda$ has joint probability density function $f;a_1,a_2$ as in (1.1) and $\mu$ has joint probability density function $f;\beta,a_1$ as in (1.2).

Inspired by the work in [9] and [11], for general $\beta > 0$, we want to completely understand the behavior between $\lambda$ and $\mu$. Making a minor adjustment from $d(\mathcal{L}(2a_2\lambda), \mathcal{L}(\mu))$ in [9], we will investigate the following object 

$$d(\mathcal{L}(2a\lambda), \mathcal{L}(\mu))$$

under the condition $a_1m = o(a_2)$ or $a_1m = O(a_2)$ with $a = a_1 + a_2$.

For two different distances mentioned above, we have the following theorem.

**Theorem 1.1.** Let $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ be random variables with density $f;\beta,a_1$ as in (1.2) and $f;\beta,a_1,a_2$ as in (1.1), respectively. Let $d(\mathcal{L}(2a\lambda), \mathcal{L}(\mu))$ be the total variation distance or the Kullback–Leibler divergence between the probability distributions of $2a\lambda$ and $\mu$. Then
(i) \( \lim_{a_2 \to \infty} d(\mathcal{L}(2a\lambda), \mathcal{L}(\mu)) = 0 \) if \( \lim_{a_2 \to \infty} \frac{a_1m}{a_2} = 0 \).

(ii) \( \lim_{a_2 \to \infty} d(\mathcal{L}(2a\lambda), \mathcal{L}(\mu)) > 0 \) if \( \lim_{a_2 \to \infty} \frac{a_1m}{a_2} = \sigma > 0 \).

By Pinsker’s inequality, we know
\[
\|\mu - \nu\|_{TV}^2 \leq 2D_{KL}(\mu\|\nu).
\]
Therefore as in [11], for the first item, we just need to prove
\[
\lim_{a_2 \to \infty} D_{KL}(\mathcal{L}(2a\lambda)\|\mathcal{L}(\mu)) = 0 \tag{1.3}
\]
and for the second item it suffices to prove
\[
\lim_{a_2 \to \infty} \|\mathcal{L}(2a\lambda) - \mathcal{L}(\mu)\|_{TV} > 0. \tag{1.4}
\]
Furthermore, for the validity of (1.4), by Lemma 2.15 in [11], it is enough to prove (1.4) under the following two conditions:

(A1) \( m \equiv 1 \) and \( \lim_{a_2 \to \infty} \frac{a_1}{a_2} \in (0, 1) \);

(A2) \( m \to \infty, \lim_{a_2 \to \infty} \frac{\beta m}{2a_1} = \gamma \in [0, 1] \) and \( \lim_{a_2 \to \infty} \frac{ma_1}{a_2} = \sigma > 0 \).

Define
\[
K_m = \left( \frac{1}{a} \right)^{ma_1} \prod_{i=0}^{m-1} \frac{\Gamma(a - \frac{\beta_i}{2})}{\Gamma(a - \frac{\gamma_i}{2})},
\]
\[
L_m(x_1, \ldots, x_m) = e^{\frac{1}{2} \sum_{i=1}^{m} x_i} \prod_{i=1}^{m} \left( 1 - \frac{x_i}{2a} \right)^{a_2 - r} I_{\{\max x_i \leq 2a\}}.
\]
We will show in the fourth section that the total variation distance can be regarded as
\[
\|\mathcal{L}(2a\lambda) - \mathcal{L}(\mu)\|_{TV} = \mathbb{E}[K_m L_m(\mu) - 1].
\]
Meanwhile for the Kullback–Leibler divergence, we understand it as
\[
D_{KL}(\mathcal{L}(2a\lambda)\|\mathcal{L}(\mu)) = \mathbb{E} \log (K_m L_m(\lambda)). \tag{1.5}
\]
To prove (1.3), with the help of expression (1.5) and Taylor’s formula for \( \log L_m \), one just needs to characterize the asymptotics of \( \log K_m \) and to have the asymptotical expression for \( \sum_{i=1}^{m} \mathbb{E} \lambda_i^k \) with \( k = 1, 2, 3 \), where \( (\lambda_1, \ldots, \lambda_m) \) have joint density function \( f_{\beta,a_1,a_2} \). According to the interpretation of Edelman and Sutton in [6] (see also [5]), \( f_{\beta,a_1,a_2} \) is also the joint density function of the eigenvalues of \( BB' \). The explicit form of \( m \) by \( m \) random matrix \( B \) will be given later in (2.1), whose elements are related to mutually independent Beta distributions. There isn’t any result on \( \sum_{i=1}^{m} \mathbb{E} \lambda_i^k \) when \( ma_1 = o(a_2) \), which
then requires tedious calculations related to Beta distribution presented in Section 2.

For the proof of (1.4), we have to establish a central limit theorem for 
\[ \log(K_m L_m(\mu)) \] as in [11]. Review \( r = 1 + \frac{\beta(m-1)}{2} \) and set

\[ U_m = \frac{r}{2a_2} \sum_{i=1}^{m} (\mu_i - 2a_1) - \frac{(a_2 - r)}{8a_2^2} \sum_{i=1}^{m} (\mu_i - 2a_1)^2. \]

With the help of Taylor’s formula and the property of logarithmic Gamma function, we are able to write

\[ \log(K_m L_m(\mu)) = U_m - \mathbb{E}U_m + C_m. \]

Here \( C_m \to c \) in probability for some constant \( c \) as \( a_2 \to \infty \) when (A2) is satisfied. Theorem 1.5 in [4] offered a central limit theorem for linear statistics of \( \mu \) under the assumption (A2) with the restriction \( \gamma > 0 \), based on which we could obtain a central limit theorem for \( U_m - c_m \). Here \( c_m \) satisfies

\[ \lim_{a_2 \to \infty} \frac{c_m}{\mathbb{E}U_m} = 1 \quad \text{and} \quad c_m - \mathbb{E}U_m = o(a_1), \]

which is not applicable under our condition. Besides, the case \( \gamma = 0 \) in the assumption (A2) is not discussed in [4]. Hence, we will use the modified version of Theorem 1.5 in [4], given in [7].

This paper will be organized as follows: the second section is devoted to the preparations, including the moments on beta-Jacobi ensembles \( \lambda \) and some estimates and limit theorems on beta-Laguerre ensembles \( \mu \); we give the proof of Theorem 1.1 in the third section.

2 Preliminaries

In this section, we collect all lemmas and propositions we need. We first give a lemma on the asymptotic of \( K_m \).

**Lemma 2.1.** For \( 0 \leq \beta(m-1) < 2(a_1 \wedge a_2) \), recall \( a = a_1 + a_2, \eta = \frac{\beta}{2} \) and

\[ K_m = \left( \frac{1}{a} \right)^{ma_1} \prod_{i=0}^{m-1} \frac{\Gamma(a - \eta i)}{\Gamma(a_2 - i\eta)} , \]

Suppose \( a_1 \to \infty \) and \( a_1 m = O(a_2) \) as \( a_2 \to \infty \). Then

\[ \log K_m = -a_1 m + m \left( a_2 - \frac{r}{2} \right) \log \left( 1 + \frac{a_1}{a_2} \right) - \frac{\beta^2 a_1 m^3}{24a_2} + o(1), \]

where \( r = 1 + \frac{\beta}{2}(m - 1) \).

The proof of this lemma is similar to that of Lemma 2.7 in [11] and is omitted here.
2.1 On \((\lambda_1, \lambda_2, \ldots, \lambda_m)\) Having Joint Density Function \(f_{\beta, a_1, a_2}\)

Now we want to understand what \(\mathbb{E} \sum_{i=1}^{m} \lambda_i^k\) for \(k = 1, 2, 3\) will be when \(a_1m = o(a_2)\). However this asymptotic could not be provided by the explicit form of the joint density (1.1). Therefore we need the help of the interpretation from Edelman and Sutton ([6], see also [5]) as mentioned in the Introduction. That is, the eigenvalues of \(BB'\) have joint density function \(f_{\beta, a_1, a_2}\), where the \(m\) by \(m\) random matrix \(B\) has the form

\[
B = \begin{pmatrix}
\sqrt{c_m s_{m-1}} & \sqrt{c_{m-1} s_{m-2}} & \cdots \\
-\sqrt{s_{m-1} c_{m-1}} & \sqrt{c_{m-2} s_{m-3}} & \cdots \\
& \ddots & \ddots \\
& & -\sqrt{s_1 c_1}
\end{pmatrix}
\]

(2.1)

with the non-negative random variables \(c_i, s_i, i = 1, 2, \ldots, m\) and \(c_i', s_i', i = 1, 2, \ldots, m - 1\) obeying the distribution and relationships

1) \(\{c_1, c_2, \ldots, c_m, c'_1, c'_2, \ldots, c'_{m-1}\}\) mutually independent;
2) \(c_i \sim \text{Beta}(a_1 - \eta(m - i), a_2 - \eta(m - i))\);
3) \(c_i' \sim \text{Beta}(\eta i, a_1 + a_2 - \eta(2m - i - 1))\);
4) \(s_i + c_i = 1, s_i' + c_i' = 1\).

Based on their interpretation, Dumitriu and Paquette [5] obtained a series expansion of the scaled moment \(\frac{1}{m} \text{Etr}((BB')^k)\) when \(a_1, a_2\) and \(m\) have same order. Precisely, \(\frac{1}{m} \text{Etr}((BB')^k) = \sum_{j=0}^{\infty} \rho_k(j, \alpha)m^{-j}\). The coefficients \(\rho_k(j, \alpha)\) are palindromic polynomials in \((-\alpha)\) of degree \(j\). This result is perfect with concise form, but the exact expression of \(\frac{1}{m} \text{Etr}((BB')^k)\) depending on \(a_1, a_2\) and \(m\) is hidden in \(\rho_k(j, \alpha)\). Therefore, with the help of (2.1), we calculate directly the following expressions under \(a_1m = o(a_2)\).

**Proposition 2.1.** Suppose that \(a_1m = o(a_2)\) as \(a_2 \to \infty\). Assume that \((\lambda_1, \lambda_2, \ldots, \lambda_m)\) have joint probability density function \(f_{\beta, a_1, a_2}\) as in (1.1). Then we have as \(a_2 \to \infty\) with \(a = a_1 + a_2\),

\[
\mathbb{E} \sum_{i=1}^{m} \lambda_i = \frac{a_1m}{a} + o(ma_2^{-1});
\]

\[
\mathbb{E} \sum_{i=1}^{m} \lambda_i^2 = \frac{1}{a^2}(a_1^2m^2 + \eta a_1 m^2) + o(a_2^{-1});
\]

\[
\mathbb{E} \sum_{i=1}^{m} \lambda_i^3 = \frac{1}{a^3}(a_1^3m^3 + 3\eta a_1^2 m^2 + \eta^2 a_1 m^3) + o(a_2^{-1}).
\]

(2.2)
Proof. By the interpretation above, with the convention $s'_0 = 1$, we have

$$
\mathbb{E} \sum_{i=1}^{m} \lambda_i = \mathbb{E} \text{tr}(BB') = \mathbb{E} \sum_{i=1}^{m} \mathbb{E} c_{m+1-i}s'_{m-i} + \mathbb{E} s_{m-i}c'_{m-i} \quad (2.3)
$$

and

$$
\mathbb{E} \sum_{i=1}^{m} \lambda_i^2 = \mathbb{E} \text{tr}((BB')^2) = \mathbb{E} \sum_{i=1}^{m-1} s^2_{m-i}(c'_{m-i})^2 + \mathbb{E} \sum_{i=1}^{m} c^2_{m+1-i}(s'_{m-i})^2 + 2\mathbb{E} \sum_{i=1}^{m-1} s_{m-i}c'_{m-i}s'_{m-i-1} + 2\mathbb{E} \sum_{i=1}^{m-1} c_{m+1-i}s_{m-i}s'_{m-i} \quad (2.4)
$$

According to the expressions (2.3) and (2.4), for $\mathbb{E} \sum_{i=1}^{m} \lambda_i$ and $\mathbb{E} \sum_{i=1}^{m} \lambda_i^2$, we have to work on the following six items:

$$
\mathbb{E}c_{m-i}, \mathbb{E}c'_{m-i}, \mathbb{E}c^2_{m-i}, \mathbb{E}(c'_{m-i})^2, \mathbb{E}s^2_{m-i} \text{ and } \mathbb{E}(s'_{m-i})^2.
$$

For the random variable $\xi \sim \text{Beta}(\alpha, \beta)$, it is well-known that

$$
\mathbb{E}\xi = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \mathbb{E}\xi^2 = \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)}. \quad (2.5)
$$

Since $2a_1 > \beta(m-1)$, it enforces that $m^2/a_2 \to 0$ when $a_2 \to \infty$ from $a_1m = o(a_2)$. By definition and (2.5), keeping in mind that $a = a_1 + a_2$, one gets

$$
\mathbb{E}c_{m-i} = \frac{a_1 - \eta i}{a - 2\eta i} = \frac{a_1 - \eta i}{a} + o(a_2^{-1}) \quad (2.6)
$$

for $0 \leq i \leq m - 1$. Here and after we use frequently the following trick to make $i$ vanish from the denominator as for (2.6). That is

$$
\frac{a_1 - \eta i}{a - 2\eta i} = \frac{a_1 - \eta i}{a} \left( \frac{a - 2\eta i + 2\eta i}{a - 2\eta i} \right) = \frac{a_1 - \eta i}{a} + \frac{2\eta i(a_1 - \eta i)}{a(a - 2\eta i)} = \frac{a_1 - \eta i}{a} + o(a_2^{-1}),
$$

where the last equality holds since $i(a_1 - \eta i) \leq a_1m = o(a_2) = o(a)$. Similarly, we have

$$
\mathbb{E}c'_{m-i} = \frac{\eta(m-i)}{a - \eta(2i-1)} = \frac{\eta(m-i)}{a} + o(a_2^{-1});
$$

$$
\mathbb{E}c^2_{m-i} = \frac{(a_1 - \eta i)(a_1 - \eta i + 1)}{(a - 2\eta i)(a - 2\eta i + 1)} = \frac{(a_1 - \eta i)^2 + (a_1 - \eta i)}{a^2} + o\left(\frac{a_1}{a^2}\right); \quad (2.7)
$$

$$
\mathbb{E}(c'_{m-i})^2 = \frac{\eta(m-i)(\eta(m-i) + 1)}{(a - 2\eta i + \eta)(a - 2\eta i + \eta + 1)} = \frac{\eta^2(m-i)^2}{a^2} + o(a_2^{-3/2}).
$$
Consequently
\[\mathbb{E} s_m^2 = \mathbb{E} (1 - c_m)^2 = 1 - \frac{2(a_1 - \eta i)}{a} + \frac{(a_1 - \eta i)^2 + (a_1 - \eta i)}{a^2} + o(a_2^{-1});\]
\[\mathbb{E} (s'_m)^2 = \mathbb{E} (1 - c_m')^2 = 1 - \frac{2\eta (m - i)}{a} + o(a_2^{-1}).\]

Plugging (2.6) and (2.7) into expression (2.3), we have
\[
\mathbb{E} \sum_{i=1}^{m} \lambda_i = \sum_{i=1}^{m} \mathbb{E} c_{m+1-i} + \sum_{i=1}^{m-1} \mathbb{E} c_{m-i} - \sum_{i=1}^{m} \mathbb{E} c_{m+1-i} \mathbb{E} c_{m-i} - \sum_{i=1}^{m-1} \mathbb{E} c_{m-i} \mathbb{E} c_{m-i}'
\]
\[= \frac{1}{a} \sum_{i=1}^{m-1} (a_1 - \eta(i-1) + \eta(m - i)) + \frac{a_1}{a} + o(ma_2^{-1})
\]
\[= \frac{a_1 m}{a} + o(ma_2^{-1}).\]

Here for the second term, we use the facts \(a_1 m = o(a_2), (2.6)\) and (2.7) to get
\[
\sum_{i=1}^{m} \mathbb{E} c_{m+1-i} \mathbb{E} c_{m-i}' = \sum_{i=1}^{m} O(a_1 ma^{-2}) = O(a_1 m^2 a_2^{-2}) = o(ma_2^{-1}).
\]

Next we focus on the second expression in (2.2). We treat the first term of (2.4). Since \(m = o(\sqrt{a_2})\) and \(o(ma_2^{-3/2}) = o(a_2^{-1})\), we can drop off the terms in the sum \(\sum_{i=1}^{m}\), which have order \(o(a_2^{-s})\) with \(s \geq 3/2\). This would greatly simplify the calculus. Thereby, based on (2.7), (2.8) and the condition \(a_1 m = o(a_2)\) and \(\beta(m - 1) < 2a_1\), it follows from the independence of \(\{c_1, \ldots, c_m, c'_1, \ldots, c'_{m-1}\}\) that
\[
\mathbb{E} \sum_{i=1}^{m-1} s_{m-i}^2 (c_{m-i}')^2 = \sum_{i=1}^{m-1} \left( \left(1 - \frac{a_1 - \eta i}{a}\right)^2 + \frac{a_1 - \eta i}{a^2} + o(a_2^{-1}) \right)
\times \left( \frac{\eta^2 (m - i)^2}{a^2} + o(a_2^{-3/2}) \right)
\]
\[= \sum_{i=1}^{m-1} \left( \frac{\eta^2 (m - i)^2}{a^2} + o(a_2^{-3/2}) \right)
\]
\[= \frac{\eta^2 m^3}{3a^2} + o(a_2^{-1}),\]

where for the second equality we drop off directly the term
\[
\left( - \frac{2(a_1 - \eta i)}{a} + \frac{(a_1 - \eta i)^2 + (a_1 - \eta i)}{a^2} + o(a_2^{-1}) \right) \frac{\eta^2 (m - i)^2}{a^2} = o(a_2^{-3/2})
\]
for any $1 \leq i \leq m - 1$. Similarly since $a_1 m = o(a_2)$, we have
\[
\mathbb{E} \sum_{i=1}^{m} c_{m+1-i}^2 (s_{m-i}^\prime)^2 = \sum_{i=1}^{m} \left( \frac{(a_1 - \eta_i + \eta)^2 (a_1 - \eta_i + \eta)}{a^2} + o \left( \frac{a_1}{a^2} \right) \right) \\
\times \left( 1 - \frac{2\eta(m-i)}{a} + o(a_2^{-1}) \right) \\
= \sum_{i=1}^{m} \left( \frac{(a_1 - \eta_i)^2}{a^2} + O \left( \frac{a_1}{a^2} \right) \right) \\
= \frac{3a_1^2}{3a^2} - 3a_1m^2 + \eta^2m^3 + o(a_2^{-1}).
\]
The same argument also leads
\[
\mathbb{E} \sum_{i=1}^{m-1} s_{m-i} c_{m-i} c_{m-i}^\prime s_{m-i}^\prime - 1 = \sum_{i=1}^{m-1} (\mathbb{E} c_{m-i} - \mathbb{E} c_{m-i}^2) \mathbb{E} c_{m-i}^\prime (1 - \mathbb{E} c_{m-i}^\prime) \\
= \sum_{i=1}^{m-1} \left( \frac{\eta(a_1 - \eta_i)(m-i)}{a^2} + o \left( \frac{a_1}{a^2} \right) \right) \\
= \frac{3\eta a_1m^2 - \eta^2m^3}{6a^2} + o(a_2^{-1})
\]
and
\[
\mathbb{E} \sum_{i=1}^{m-1} c_{m+1-i} c_{m-i}^\prime s_{m-i} s_{m-i}^\prime = \sum_{i=1}^{m-1} \left( \frac{\eta(a_1 - \eta_i)(m-i)}{a^2} + o \left( \frac{a_1}{a^2} \right) \right) \\
= \frac{3\eta a_1m^2 - \eta^2m^3}{6a^2} + o(a_2^{-1}).
\]
Therefore plugging all these four expressions above into (2.4), we have
\[
\mathbb{E} \sum_{i=1}^{m} \lambda_i^2 = \frac{\eta^2m^3}{3a^2} + \frac{3a_1^2}{3a^2} - 3\eta a_1m^2 + \eta^2m^3 + 4 \cdot \frac{3\eta a_1m^2 - \eta^2m^3}{6a^2} + o(a_2^{-1}) \\
= \frac{a_1^2m + \eta a_1m^2}{a^2} + o(a_2^{-1}).
\]
Now we work on the last expression in (2.2). For the Beta distribution $\xi \sim \text{Beta}(\alpha, \beta)$, one knows
\[
\mathbb{E} \xi^3 = \frac{\alpha(\alpha + 1)(\alpha + 2)}{(\alpha + \beta)(\alpha + \beta + 1)(\alpha + \beta + 2)}.
\]
Therefore we have
\[
\mathbb{E} c_{m-i}^3 = \frac{(a_1 - \eta_i)^3}{a^3} + o(a_1 a_2^{-2});
\]
\[
\mathbb{E} (c_{m-i}^\prime)^3 = \frac{\eta^3(m-i)^3}{a^3} + o(a_1 a_2^{-2}).
\]
With careful calculation, we have

\[
\mathbb{E} \sum_{i=1}^{m} \lambda_i^3 = \sum_{i=1}^{m} \mathbb{E} c_{m-i+1}^3 (1 - \mathbb{E} c_{m-i}'^3) + \sum_{i=1}^{m-1} \mathbb{E} (c_{m-i}')^3 (1 - \mathbb{E} c_{m-i})^3 \\
+ 3 \sum_{i=1}^{m-1} (1 - \mathbb{E} c_{m-i})^2 (\mathbb{E} (c_{m-i}')^2 - \mathbb{E} (c_{m-i})^3) \mathbb{E} c_{m-i+1} \\
+ 3 \sum_{i=1}^{m-1} (1 - \mathbb{E} c_{m-i-1})^2 (\mathbb{E} c_{m-i}^2 - \mathbb{E} c_{m-i}') \mathbb{E} c_{m-i} \\
+ 3 \sum_{i=1}^{m-1} (1 - \mathbb{E} c_{m-i}) \mathbb{E} (c_{m-i}' (1 - c_{m-i})^2) \mathbb{E} c_{m-i+1}^2 \\
+ 3 \sum_{i=1}^{m-1} (1 - \mathbb{E} c_{m-i-1}) \mathbb{E} (c_{m-i} (1 - c_{m-i})^2) \mathbb{E} (c_{m-i}')^2 \\
+ 3 \sum_{i=1}^{m-1} \mathbb{E} (c_{m-i} - c_{m-i}') \mathbb{E} (c_{m-i}' - (c_{m-i}')^2) \mathbb{E} c_{m-i+1} (1 - \mathbb{E} c_{m-i-1}) \\
+ 3 \sum_{i=1}^{m-1} \mathbb{E} (c_{m-i+1} - c_{m-i}'^3) \mathbb{E} (c_{m-i}' - (c_{m-i})^2) \mathbb{E} c_{m-i+1} (1 - \mathbb{E} c_{m-i}).
\]

Similarly as for \( \mathbb{E} \sum_{i=1}^{m} \lambda_i^2 \), we drop off the terms \( o(a_1 a_2^{-2}) \). Plugging (2.6), (2.7) and (2.9) into above expression, one gets

\[
\mathbb{E} \sum_{i=1}^{m} \lambda_i^3 = \sum_{i=1}^{m} \frac{(a_1 - \eta)^3}{a^3} + \sum_{i=1}^{m-1} \frac{\eta^3 (m - i)^3}{a^3} + o(a_2^{-1}) \\
+ 9 \frac{m-1}{a^3} \sum_{i=1}^{m-1} \{ \eta (m - i) (a_1 - \eta i)^2 + \eta^2 (m - i)^2 (a_1 - \eta i) \} \\
= \frac{1}{a^3} (a_1^3 m + 3 \eta a_1^2 m^2 + \eta^2 a_1 m^3) + o(a_2^{-1}).
\]

This finally closes the entire proof. \( \Box \)

2.2 On \( \mu = (\mu_1, \mu_2, \ldots, \mu_m) \) Having Density Function \( f_{\beta,a_1} \)

According to the famous characterization of Dumitriu and Edelman in [3] and [2], we know that \( (\mu_i)_{1 \leq i \leq m} \) are the eigenvalues of the matrix \( AA' \), where \( A \) is
given as

\[
A = \begin{pmatrix}
x_1 & x_2 \\
y_2 & x_3 \\
\vdots & \ddots \\
y_m & x_m
\end{pmatrix}
\] (2.10)

with the non-negative random variables \(x_i, i = 1, 2, \ldots, m\) and \(y_i, i = 2, \ldots, m\) obeying the distribution and relationships

1) \(\{x_1, x_2, \ldots, x_m, y_2, y_3, \ldots, y_m\}\) mutually independent;

2) \(x_i^2 \sim \chi^2_{(2a_1 - \beta(i-1))}\) and \(y_i^2 \sim \chi^2_{\beta(m-(i-1))}\).

Since our calculus below will heavily depend on the properties of \(\chi^2\)-distribution, we present Lemma 2.8 in [11].

**Lemma 2.2.** Given a random variable \(X\) which obeys the \(\chi^2\)-distribution with parameter \(n\) for some \(n \geq 1\), i.e., \(X \sim \chi^2_n\), then we have

\[
\mathbb{E} X^k = \prod_{l=0}^{k-1} (n + 2l), \quad \forall k \geq 1;
\]

\[
\mathbb{E} (X - n)^2 = 2n, \quad \mathbb{E} (X - n)^3 = 8n;
\]

\[
\mathbb{E} (X - n)^4 = 12n(n + 4), \quad \text{Var}(X^2) = 8n(n + 2)(n + 3);
\]

\[
\text{Var}((X - n)^2) = 8n(n + 6).
\]

Now we present two key lemmas, whose proofs are relatively long and will be postponed to the appendix.

**Lemma 2.3.** Let \(\mu = (\mu_1, \mu_2, \ldots, \mu_m)\) be the random variables having joint distribution density \(f_{\beta, a_1}\) given in (1.2). We have

\[
\text{Var} \left( \sum_{i=1}^{m} \mu_i \right) = 4a_1m, \quad \mathbb{E} \sum_{i=1}^{m} (\mu_i - 2a_1)^2 = 4a_1mr; \]

\[
\text{Var} \left( \sum_{j=1}^{m} (\mu_j - 2a_1)^2 \right) = 16\beta a_1^2 m^2 + 16\beta^2 a_1^3 m^3 + o(a_1^2 m^2); \]

\[
\text{Cov} \left( \sum_{i=1}^{m} (\mu_i - 2a_1), \sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right) = 16a_1mr; \]

\[
\mathbb{E} \sum_{i=1}^{m} (\mu_i - 2a_1)^3 = 2\beta^2 a_1 m^3 + o(a_1 m^3)
\]

for \(m \geq 2\) and \(a_1 > \frac{\beta}{2}(m - 1)\).
Lemma 2.4. Let \( \mu = (\mu_1, \mu_2, \ldots, \mu_m) \) be the random variables having joint distribution density \( f_{\beta, a_1} \) given in (1.2). Suppose that \( a_1 = O(m) \), then we have
\[
\text{Var}\left( \sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right) = O(m^7)
\]
for \( m \) sufficiently large.

Next we give a lemma to describe the property of \( \max_{1 \leq i \leq m} |\mu_i - 2a_1| \) under the assumption (A2).

Lemma 2.5. Let \( (\mu_1, \mu_2, \ldots, \mu_m) \) be the random variables having joint distribution density \( f_{\beta, a_1} \) given in (1.2). Suppose the assumption (A2) holds, i.e.,
\[
\lim_{a_2 \to \infty} \frac{a_1 m}{a_2} = \sigma > 0 \quad \text{and} \quad \lim_{a_1 \to \infty} \frac{m}{a_1} = \gamma \in [0, 1].
\]
Then we have
\[
\max_{1 \leq i \leq m} \left| \frac{\mu_i}{2a_1} - 1 \right| \leq (\sqrt{\gamma} + 1)^2
\]
with probability one and
\[
\max_{1 \leq i \leq m} \left| \frac{\mu_i - 2a_1}{2a_2} \right| \xrightarrow{p} 0
\]
as \( a_2 \to \infty \).

Proof. Set \( \mu_{\text{max}} = \max_{1 \leq i \leq m} \mu_i \) and \( \mu_{\text{min}} = \min_{1 \leq i \leq m} \mu_i \). Theorem 1.1 in [7] tells
\[
\frac{\mu_{\text{max}}}{2a_1} \to (1 + \sqrt{\gamma})^2 \quad \text{and} \quad \frac{\mu_{\text{min}}}{2a_1} \to (1 - \sqrt{\gamma})^2
\]
in probability as \( m \to \infty \). Since for any \( 1 \leq i \leq m \)
\[
(1 - \sqrt{\gamma})^2 - 2 \leq \frac{\mu_{\text{min}}}{2a_1} - 1 \leq \frac{\mu_i}{2a_1} - 1 \leq \frac{\mu_{\text{max}}}{2a_1} - 1 \leq (\sqrt{\gamma} + 1)^2,
\]
which implies with probability one that as \( m \) large enough
\[
\max_{1 \leq i \leq m} \left| \frac{\mu_i}{2a_1} - 1 \right| \leq (1 + \sqrt{\gamma})^2.
\]
The proof is then completed since
\[
\max_{1 \leq i \leq m} \left| \frac{2a_1 - \mu_i}{2a_2} \right| = \frac{2a_1}{2a_2} \max_{1 \leq i \leq m} \left| \frac{\mu_i - 2a_1}{2a_1} \right| \leq (1 + \sqrt{\gamma})^2 \frac{a_1}{a_2}.
\]
3 Proof of Theorem 1.1

In this section, we will give the final statement on the proof of Theorem 1.1.

Recall the joint density function of \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in [0, 1]^m \) is given by

\[
f_{\beta, a_1, a_2}(x_1, x_2, \ldots, x_m) = C_{J}^{\beta, a_1, a_2} \prod_{1 \leq i < j \leq m} |x_i - x_j|^\beta \prod_{i=1}^{m} x_i^{a_1-r}(1-x_i)^{a_2-r}, \tag{3.1}
\]

where \( a_1, a_2 > \frac{\beta(m-1)}{2} \) and \( r = 1 + \frac{\beta}{2}(m - 1) \), and

\[
C_{J}^{\beta, a_1, a_2} = \prod_{j=1}^{m} \frac{\Gamma(1 + \beta/(2j)) \Gamma(a_1 + a_2 - (\beta/(2))(m-j))}{\Gamma(1 + \beta/(2)) \Gamma(a_1 - (\beta/(2))(m-j)) \Gamma(a_2 - (\beta/(2))(m-j))}.
\]

It is clear that the joint distribution density for \( \theta := 2a \lambda \), denoted by \( g_{\beta, a_1, a_2} \), should be as follows

\[
g_{\beta, a_1, a_2}(x_1, x_2, \ldots, x_m) := C_{J}^{\beta, a_1, a_2} \left( \frac{1}{2a} \right)^c \prod_{1 \leq i < j \leq m} |x_i - x_j|^\beta \prod_{i=1}^{m} x_i^{a_1-r}(1-x_i)^{a_2-r} I_{\{\max \theta_i \leq 2a\}} \tag{3.2}
\]

where \( c := \beta m(m - 1)/2 + m(a_1 - r) + m = a_1 m \). Review the joint density function of \( \mu = (\mu_1, \mu_2, \ldots, \mu_m) \) is

\[
f_{\beta, a_1}(x_1, x_2, \ldots, x_m) = C_{L}^{\beta, a_1} \prod_{1 \leq i < j \leq m} |x_i - x_j|^\beta \prod_{i=1}^{m} x_i^{a_1-r} e^{-1/2 \sum_{i=1}^{m} x_i} \tag{3.3}
\]

where

\[
C_{L}^{\beta, a_1} = 2^{-ma_1} \prod_{j=1}^{m} \frac{\Gamma(1 + \beta/(2j)) \Gamma(a_1 - (\beta/(2))(m-j))}{\Gamma(1 + \beta/(2)) \Gamma(a_1 - (\beta/(2))(m-j)))}.
\]

Remember

\[
K_m = \left( \frac{1}{a} \right)^{ma_1} \prod_{i=0}^{m-1} \frac{\Gamma(a - \eta i)}{\Gamma(a_2 - \eta i)}; \tag{3.3}
\]

\[
L_m(x_1, \ldots, x_m) = e^{\frac{1}{2} \sum_{i=1}^{m} x_i} \prod_{i=1}^{m} \left( 1 - \frac{x_i}{2a} \right)^{a_2-r} I_{\{\max x_i \leq 2a\}}.
\]

Observing the expressions (3.1), (3.2) and (3.3), we know

\[
\frac{g_{\beta, a_1, a_2}}{f_{\beta, a_1}} = K_m L_m(x_1, \ldots, x_m).
\]
This leads
\[ \| \mathcal{L}(2a\lambda) - \mathcal{L}(\mu) \|_{TV} = \int_{[0, \infty)^m} |g_{\beta,a_1,a_2}(x) - f_{\beta,a_1}(x)| dx \]
\[ = \int_{[0, \infty)^m} \left| \frac{g_{\beta,a_1,a_2}(x)}{f_{\beta,a_1}(x)} - 1 \right| f_{\beta,a_1}(x) dx \]
\[ = \mathbb{E} |K_m L_m(\mu) - 1|. \]

Meanwhile, the Kullback–Leibler divergence \( D_{KL}(\mathcal{L}(2a\lambda)||\mathcal{L}(\mu)) \) could be expressed as
\[ D_{KL}(\mathcal{L}(2a\lambda)||\mathcal{L}(\mu)) = \int_{[0, \infty)^m} \frac{g_{\beta,a_1,a_2}(x)}{f_{\beta,a_1}(x)} \log \frac{g_{\beta,a_1,a_2}(x)}{f_{\beta,a_1}(x)} f_{\beta,a_1}(x) dx \]
\[ = \int_{[0, \infty)^m} \log \frac{g_{\beta,a_1,a_2}(x)}{f_{\beta,a_1}(x)} g_{\beta,a_1,a_2}(x) dx \]
\[ = \mathbb{E} \log (K_m L_m(\lambda)). \]

As in [11], we consider another modified version \( L'_m, K'_m \) of \( L_m, K_m \) respectively, which are defined by
\[ L'_m = \left(1 + \frac{a_1}{a_2}\right)^{m(a_2-r)} L_m; \]
\[ K'_m = \left(1 + \frac{a_1}{a_2}\right)^{-m(a_2-r)} K_m. \] (3.4)

Obviously \( L'_m K'_m = L_m K_m \). Therefore we have
\[ \| \mathcal{L}(2a\lambda) - \mathcal{L}(\mu) \|_{TV} = \mathbb{E} |K'_m L'_m(\mu) - 1|; \]
\[ D_{KL}(\mathcal{L}(2a\lambda)||\mathcal{L}(\mu)) = \mathbb{E} \log (K'_m L'_m(\lambda)). \] (3.5)

### 3.1 Proof of (i) of Theorem 1.1

As mentioned in the introduction, to prove (i) of Theorem 1.1, we just need to prove
\[ \lim_{a_2 \to \infty} D_{KL}(\mathcal{L}(2a\lambda)||\mathcal{L}(\mu)) = 0. \]

By Lemma 2.1, since \( a_1 m = o(a_2) \) and \( \beta(m - 1) < 2a_1 \), then one gets \( a_1 m^3 = o(a_2^2) \). Recalling \( r = \frac{\beta}{2}(m - 1) + 1 \), we have
\[ \log K'_m = \log (K_m) - m(a_2-r) \log \left(1 + \frac{a_1}{a_2}\right) \]
\[ = -a_1 m + \frac{rm}{2} \log \left(1 + \frac{a_1}{a_2}\right) + o(1) \]
\[ = -a_1 m + \frac{\eta a_1 m^2}{2a_2} + o(1). \]
Meanwhile by (3.3) and (3.4), we have

\[ \mathbb{E} \log(L'_m(\lambda)) = m(a_2 - r) \log \left(1 + \frac{a_1}{a_2}\right) + \frac{1}{2} \mathbb{E} \sum_{i=1}^{m} \theta_i \]

\[ + (a_2 - r) \mathbb{E} \sum_{i=1}^{m} \log \left(1 - \frac{\theta_i}{2a}\right) \]  \hspace{1cm} (3.6)

\[ = \frac{1}{2} \mathbb{E} \sum_{i=1}^{m} \theta_i + (a_2 - r) \mathbb{E} \sum_{i=1}^{m} \log \left(1 + \frac{2a_1 - \theta_i}{2a_2}\right). \]

Here we use the fact

\[ \log \left(1 + \frac{a_1}{a_2}\right) + \log \left(1 - \frac{\theta_i}{2a}\right) = \log \left(1 + \frac{2a_1 - \theta_i}{2a_2}\right). \]

Therefore it follows from (3.5) that

\[ D_{KL}(\mathcal{L}(2a\lambda)||\mathcal{L}(\mu)) = \mathbb{E} \log K'_m + \mathbb{E} \log(L'_m(\lambda)) \]

\[ = -a_1m + \frac{\eta a_1 m^2}{2a_2} + \frac{1}{2} \mathbb{E} \sum_{i=1}^{m} \theta_i + (a_2 - r) \mathbb{E} \sum_{i=1}^{m} \log \left(1 + \frac{2a_1 - \theta_i}{2a_2}\right) + o(1) \]

\[ \leq -a_1m + \frac{\eta a_1 m^2}{2a_2} + \frac{1}{2} \mathbb{E} \sum_{i=1}^{m} \theta_i + o(1) \]

\[ + (a_2 - r) \mathbb{E} \sum_{i=1}^{m} \left(\frac{2a_1 - \theta_i}{2a_2} - \frac{(2a_1 - \theta_i)^2}{8a_2^2} + \frac{(2a_1 - \theta_i)^3}{24a_2^3}\right), \]

(3.7)

where the last inequality is due to the elementary inequality

\[ \log(1 + x) \leq x - \frac{x^2}{2} + \frac{x^3}{3}, \quad x > -1. \]

Obviously, applying Proposition 2.1 to \( \frac{\theta}{2a} \), we have by \( a_1 m = o(a_2) \) that

\[ -a_1m + \frac{1}{2} \mathbb{E} \sum_{i=1}^{m} \theta_i + \frac{a_2 - r}{2a_2} \mathbb{E} \sum_{i=1}^{m} (2a_1 - \theta_i) \]

\[ = -a_1m + \frac{r}{2a_2} \mathbb{E} \sum_{i=1}^{m} \theta_i + \frac{(a_2 - r)a_1 m}{a_2} \]

\[ = -a_1m + \frac{(a_2 - r)a_1 m}{a_2} + \frac{r}{a_2} a_1 m + o(1) \]

\[ = o(1) \]
and also
\[
\frac{a_2 - r}{8a_2^2} \mathbb{E} \sum_{i=1}^m (2a_1 - \theta_i)^2 = \frac{a_2 - r}{8a_2^2} \left( 4a_1^2 m - 4a_1 \mathbb{E} \sum_{i=1}^m \theta_i + \mathbb{E} \sum_{i=1}^m \theta_i^2 \right)
\]
\[
= \frac{a_2 - r}{8a_2^2} (4a_1^2 m - 4a_1 \cdot 2a_1 m + 4a_1^2 m + 4\eta a_1 m^2) + o(1)
\]
\[
= \frac{\eta a_1 m^2}{2a_2} + o(1).
\]

Similarly we get
\[
\frac{a_2 - r}{24a_2^3} \mathbb{E} \sum_{i=1}^m (2a_1 - \theta_i)^3 = \frac{a_2 - r}{24a_2^3} \mathbb{E} \sum_{i=1}^m (8a_1^3 - 12a_2^2 \theta_i + 6a_1 \theta_i^2 - \theta_i^3)
\]
\[
= \frac{a_2 - r}{24a_2^3} \left( 8a_1^3 m - 12a_1^2 \cdot 2a_1 m + 6a_1(4a_1^2 m + 4\eta a_1 m^2) - 8a_1^2 m - 24\eta a_1^2 m^2 - 8\eta^2 a_1 m^3 \right) + o(1)
\]
\[
= - \frac{(a_2 - r)\eta^2 a_1 m^3}{3a_2^3} + o(1) = o(1).
\]

Therefore plugging all these expressions into (3.7), we have

\[
D_{KL}(\mathcal{L}(2a\lambda)||\mathcal{L}(\mu)) \leq -a_1 m + \frac{\eta a_1 m^2}{2a_2} + \frac{1}{2} \mathbb{E} \sum_{i=1}^m \theta_i
\]
\[
+ (a_2 - r) \mathbb{E} \sum_{i=1}^m \left( \frac{2a_1 - \theta_i}{2a_2} - \frac{(2a_1 - \theta_i)^2}{8a_2^2} + \frac{(2a_1 - \theta_i)^3}{24a_2^3} \right)
\]
\[
= \frac{\eta a_1 m^2}{2a_2} - \frac{\eta a_1 m^2}{2a_2} + o(1) = o(1).
\]

The desired result is obtained.

\[\Box\]

### 3.2 Proof of (ii) of Theorem 1.1

In this subsection, we present a crucial lemma on the central limit theorem for \(\log(L'_m(\mu))\) with \(\mu\) having probability density function \(f_{\beta,a_1}\) in (1.2). We first state a key lemma.

**Lemma 3.1.** Recall

\[
U_m := \frac{r}{2a_2} \sum_{i=1}^m (\mu_i - 2a_1) - \frac{(a_2 - r)}{8a_2^2} \sum_{i=1}^m (\mu_i - 2a_1)^2.
\]
Lemma 3.2. Let \(\mu\) be the distribution density \(f_{\beta,a_1}\) given in (1.2) and \(L_m' \) be given in (3.4). Then under the assumption (A2), we have

\[
U_m + \frac{(a_2 - r)a_1mr}{2a_2^2} \xrightarrow{w} N \left( 0, \frac{\beta \sigma^2}{4} \right),
\]

\[
\frac{(a_2 - r)}{8a_2^2} \left( \sum_{i=1}^{m} (\mu_i - 2a_1)^2 - \mathbb{E} \sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right) \xrightarrow{w} N \left( 0, \frac{\beta \sigma^2}{4} (1 + 2\gamma) \right)
\]
as \(a_2 \to \infty\).

Proof. By Theorem 1.2 in [7], as a linear function of \(\sum_{i=1}^{m} \mu_i\) and \(\sum_{i=1}^{n} \mu_i^2\), \(U_m - \mathbb{E} U_m\) converges weakly to a normal distribution \(N(0, \sigma^2)\) as \(a_2 \to \infty\) when \(\lim_{a_1 \to 0} \frac{\beta m}{2a_1} = \gamma \in [0, 1]\). Here \(\sigma^2 := \lim_{a_1 \to \infty} \text{Var}(U_m)\). By Lemma 2.3, we have \(\mathbb{E} U_m = -\frac{(a_2 - r)a_1mr}{2a_2^2}\) and

\[
\text{Var}(U_m) = \frac{r^2}{4a_2^2} \text{Var} \left( \sum_{i=1}^{m} \mu_i \right) + \frac{(a_2 - r)^2}{64a_2^4} \text{Var} \left( \sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right) - \frac{r(a_2 - r)}{8a_2^3} \text{Cov} \left( \sum_{i=1}^{m} (\mu_i - 2a_1), \sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right)
\]

\[
= \frac{\beta^2 a_1 m^3}{4a_2^2} + \frac{\beta a_1^2 m^2 + \beta^2 a_1 m^3}{4a_2^2} - \frac{\beta^2 a_1 m^3}{2a_2^2} + o \left( \frac{\beta a_1^2 m^2}{a_2^2} \right) + \frac{\beta a_1^2 m^2}{4a_2^2} + o \left( \frac{\beta a_1^2 m^2}{a_2^2} \right).
\]

Similarly, for the limit

\[
\frac{(a_2 - r)}{8a_2^2} \left( \sum_{i=1}^{m} (\mu_i - 2a_1)^2 - \mathbb{E} \sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right) \xrightarrow{w} N \left( 0, \frac{\beta \sigma^2 (1 + 2\gamma)}{4} \right),
\]

it remains to prove that

\[
\lim_{a_2 \to \infty} \text{Var} \left( \frac{(a_2 - r)}{8a_2^2} \sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right) = \frac{\beta \sigma^2 (1 + 2\gamma)}{4}.
\]

This is guaranteed by Lemma 2.3. The proof is completed now. \(\square\)

### 3.2.1 Central Limit Theorem for \(\log(L_m'(\mu))\)

**Lemma 3.2.** Let \((\mu_1, \mu_2, \ldots, \mu_m)\) be the random variables having joint distribution density \(f_{\beta,a_1}\) given in (1.2) and \(L_m'\) be given in (3.4). Then under the assumption (A2), we have

\[
\log L_m'(\mu) - a_1 m + \frac{a_1 m(a_2 - r)}{2a_2^2} \xrightarrow{w} N \left( -\frac{\beta \gamma \sigma^2}{6}, \frac{\beta \sigma^2}{4} \right)
\]

weakly as \(a_2 \to \infty\).
Proof. Applying the Taylor formula to get \( \log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} + x^4h(x) \) with \( h \) a continuous function on \((-1, +\infty)\), we are able to write

\[
\log L'_{m}(\mu) = \frac{1}{2} \sum_{i=1}^{m} \mu_i + (a_2 - r) \sum_{i=1}^{m} \left( \frac{2a_1 - \mu_i}{2a_2} - \frac{(2a_1 - \mu_i)^2}{8a_2^2} \right) + \frac{(a_2 - r)}{24a_2^3} \sum_{i=1}^{m} (2a_1 - \mu_i)^3 + (a_2 - r) \sum_{i=1}^{m} \left( \frac{2a_1 - \mu_i}{2a_2} \right)^4 h \left( \frac{2a_1 - \mu_i}{2a_2} \right).
\]

Therefore with simple algebra, we have

\[
\log L'_{m}(\mu) = a_1 m + U_m + \frac{(a_2 - r) a_1 m r}{2a_2^2} \sum_{i=1}^{m} (2a_1 - \mu_i)^3 + (a_2 - r) \sum_{i=1}^{m} \left( \frac{2a_1 - \mu_i}{2a_2} \right)^4 h \left( \frac{2a_1 - \mu_i}{2a_2} \right).
\]

It follows

\[
\log L'_{m}(\mu) - a_1 m + \frac{(a_2 - r) a_1 m r}{2a_2^2} = U_m + \frac{(a_2 - r) a_1 m r}{2a_2^2} - \frac{(a_2 - r)}{24a_2^3} \sum_{i=1}^{m} (\mu_i - 2a_1)^3 + (a_2 - r) \sum_{i=1}^{m} \left( \frac{2a_1 - \mu_i}{2a_2} \right)^4 h \left( \frac{2a_1 - \mu_i}{2a_2} \right).
\]

By Lemma 3.1 again, one gets that

\[
U_m + \frac{(a_2 - r) a_1 m r}{2a_2^2} \rightarrow N \left( 0, \frac{\beta \sigma^2}{4} \right)
\]

weakly as \( a_2 \rightarrow \infty \). By Lemmas 2.3 and 2.4, we know

\[
\frac{(a_2 - r)^2}{a_2^6} \operatorname{Var} \left( \sum_{i=1}^{m} (\mu_i - 2a_1)^3 \right) = O \left( \frac{m^7}{a_2^4} \right) \rightarrow 0
\]

and

\[
\frac{(a_2 - r)}{24a_2^3} \mathbb{E} \sum_{i=1}^{m} (\mu_i - 2a_1)^3 = \frac{(a_2 - r)}{24a_2^3} (2\beta^2 a_1 m^3 + o(a_1 m^3)) \rightarrow \frac{\beta \gamma \sigma^2}{6}
\]
as \( a_2 \to \infty \). Consequently, it follows

\[
\frac{(a_2 - r)}{24a_2^3} \sum_{i=1}^{m}(\mu_i - 2a_1)^3 \to \frac{\beta \gamma \sigma^2}{6}
\]

in probability as \( a_2 \to \infty \). Therefore to prove (3.8), it remains to prove

\[
\bar{\delta}_m := (a_2 - r) \sum_{i=1}^{m}\left(\frac{2a_1 - \mu_i}{2a_2}\right)^4 h\left(\frac{2a_1 - \mu_i}{2a_2}\right) \to 0 \tag{3.9}
\]

in probability as \( a_2 \to \infty \). By Lemma 2.5,

\[
\max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| \to 0
\]

in probability as \( a_2 \to \infty \). Since \( h \) is continuous, \( \tau := \sup_{|x| \leq 1/2} |h(x)| < \infty \). Hence, it follows

\[
P(|\bar{\delta}_m| > \epsilon) = P\left(|\bar{\delta}_m| > \epsilon, \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| \leq \frac{1}{2}\right) + P\left(|\bar{\delta}_m| > \epsilon, \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| > \frac{1}{2}\right) \leq P\left(|\bar{\delta}_m| > \epsilon, \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| \leq \frac{1}{2}\right) + P\left(\max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| > \frac{1}{2}\right)
\]

as \( a_2 \) is sufficiently large. Then under \( \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| \leq \frac{1}{2} \),

\[
|\bar{\delta}_m| \leq \tau \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right|^2 \cdot \frac{(a_2 - r)}{4a_2^2} \sum_{i=1}^{m}(\mu_i - 2a_1)^2
\]

\[
= 2\tau \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right|^2 \cdot \frac{(a_2 - r)}{8a_2^2} \left(\sum_{i=1}^{m}(\mu_i - 2a_1)^2 - 4a_1 m r\right) + \frac{\tau a_1^3 m r (a_2 - r)}{4a_2^4} \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_1}\right|^2.
\]

The upper bound converges to zero in probability because \( \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| \to 0 \) in probability and \( \max_{1 \leq i \leq m}\left|\frac{2a_1 - \mu_i}{2a_2}\right| \) is bounded with probability one by Lemma 2.5,

\[
\frac{(a_2 - r)}{8a_2^2} \left(\sum_{i=1}^{m}(\mu_i - 2a_1)^2 - 4a_1 m r\right) \to N\left(0, \frac{\beta \sigma^2(1 + 2\gamma)}{4}\right)
\]

weakly by Lemma 3.1 and \( a_1^3 m r (a_2 - r) \to 0 \) as \( a_2 \to \infty \). This, with (3.10) and Lemma 2.5 again concludes (3.9). The proof is completed. \( \Box \)
Now we are at the position to post the proof of (ii) of Theorem 1.1. It suffices to prove that
\[
\lim_{a_2 \to \infty} \| \mathcal{L}(2a\lambda) - \mathcal{L}(\mu) \|_{TV} > 0.
\] (3.11)

By Lemma 2.14 in [11], we just need to prove (3.11) under the assumptions (A2) or (A1).

3.2.2 Proof of (ii) of Theorem 1.1 Under the Assumption (A2)

Review (3.5):
\[
\| \mathcal{L}(2a\lambda) - \mathcal{L}(\mu) \|_{TV} = E|K'_m L'_m(\mu) - 1|.
\] (3.12)

Since under (A2),
\[
a_1 m^3/a^2 = a_1 m^3/a_2^2 + o(1),
\]
use Lemma 2.1 to see under (A2),
\[
\log K'_m = -a_1 m + \frac{mr}{2} \log \left(1 + \frac{a_1}{a_2}\right) - \frac{\beta^2 a_1 m^3}{24 a_2^2} + o(1)
\]
for \(a_2\) large enough. This implies
\[
\log(K'_m L'_m(\mu)) = \log L'_m(\mu) - a_1 m + \frac{(a_2 - r) a_1 m r}{2 a_2^2} - \frac{(a_2 - r) a_1 m r}{2 a_2^2} + \frac{mr}{2} \log \left(1 + \frac{a_1}{a_2}\right) - \frac{\beta^2 a_1 m^3}{24 a_2^2} + o(1)
\]
for \(a_2\) sufficiently large. Taylor's formula allows us to write
\[
\log \left(1 + \frac{a_1}{a_2}\right) = \frac{a_1}{a_2} - \frac{a_1^2}{2 a_2^2} + o \left(\frac{a_1^2}{a_2^2}\right),
\]
which ensures
\[
s_m : = -\frac{(a_2 - r) a_1 m r}{2 a_2^2} + \frac{mr}{2} \log \left(1 + \frac{a_1}{a_2}\right) - \frac{\beta^2 a_1 m^3}{24 a_2^2}
\]
\[
= \frac{a_1 m r^2}{2 a_2^2} - \frac{a_1^2 m r}{4 a_2^2} - \frac{\beta^2 a_1 m^3}{24 a_2^2} + o(1)
\]
\[
= \frac{a_1 m (\beta^2 m^2 + O(m))}{8 a_2^2} - \frac{a_1^2 m (\beta m + 2 - \beta)}{8 a_2^2} - \frac{\beta^2 a_1 m^3}{24 a_2^2} + o(1)
\]
\[
= \frac{\beta^2 a_1 m^3}{12 a_2^2} - \frac{\beta a_1^2 m^2}{8 a_2^2} + O(a_1 m^2 + a_1^2 m) + o(1).
\]

Under the assumption (A2), we know
\[
\frac{\beta^2 a_1 m^3}{12 a_2^2} - \frac{\beta a_1^2 m^2}{8 a_2^2} \to \frac{\beta \gamma \sigma^2}{6} - \frac{\beta \sigma^2}{8} \quad \text{and} \quad \frac{O(a_1 m^2 + a_1^2 m)}{a_2^2} \to 0
\]
as \(a_2 \to \infty\). Putting this back to (3.13), we have

\[
\log(K_m' L_m'(\mu)) = \log L_m'(\mu) - a_1 m + \frac{(a_2 - r)a_1 mr}{2a_2^2} + \frac{\beta \gamma \sigma^2}{6} - \frac{\beta \sigma^2}{8} + o(1).
\]

It follows from Lemma 3.2 that

\[
\log(L_m' K_m'(\mu)) \to N\left(-\frac{\beta \sigma^2}{8}, \frac{\beta \sigma^2}{4}\right)
\]

weakly as \(a_2 \to \infty\). This is equivalent to saying \(K_m' L_m'(\mu)\) converges weakly to \(e^\xi\), where \(\xi \sim N\left(-\frac{\beta \sigma^2}{8}, \frac{\beta \sigma^2}{4}\right)\). By (3.12) and the Fatou Lemma, we have

\[
\lim_{a_2 \to \infty} \|\mathcal{L}(2a\lambda) - \mathcal{L}(\mu)\|_{TV} \geq \mathbb{E}\|e^\xi - 1\| > 0.
\]

The proof is finished. \(\square\)

### 3.2.3 Proof of (ii) of Theorem 1.1 Under the Assumption (A1)

Under this particular case, we know \(r = m = 1\) and \(\frac{a_1}{a_2} \to \sigma \in (0, 1)\). Therefore Lemma 2.1 tells

\[
\log K_1' = -a_1 + \frac{1}{2} \log(1 + \sigma) + o(1).
\]

By Taylor’s expansion

\[
\log L_1'(\mu) = a_1 + \frac{\mu - 2a_1}{2a_2} - \frac{a_2 - 1}{8a_2^2}(\mu - 2a_1)^2 + (a_2 - 1)\frac{(2a_1 - \mu)^3}{a_2^3} h\left(\frac{\mu - 2a_1}{2a_2}\right)
\]

with \(h\) a continuous function on \((-1, \infty)\). Here \(\mu\) has density function \(f_{\beta, a_1}\) with \(m = r = 1\). Then

\[
\log(K_1' L_1'(\mu)) = \frac{1}{2} \log(1 + \sigma) + \frac{\mu - 2a_1}{2a_2} - \frac{a_2 - 1}{8a_2^2}(\mu - 2a_1)^2
\]

\[
+ (a_2 - 1)\frac{(2a_1 - \mu)^3}{a_2^3} h\left(\frac{\mu - 2a_1}{2a_2}\right) + o(1).
\]

Examining the form \(f_{\beta, a_1}\) in this particular case, we see \(\mu \sim \chi^2_{2a_1}\). This allows us to rewrite \(\mu - 2a_1\) as \(\mu - 2a_1 = \sum_{i=1}^{2a_1}(\xi_i^2 - 1)\) with \(\xi_i \sim N(0, 1)\) for \(1 \leq i \leq 2a_1\) and \((\xi_i)_{1 \leq i \leq 2a_1}\) are mutually independent. Since \(\mathbb{E}\xi_i^2 = 1\) and \(\text{Var}(\xi_i^2 - 1) = 2\), by Lindeberg–Lévy central limit Theorem, we see

\[
\frac{\mu - 2a_1}{2\sqrt{a_1}} \to N(0, 1)
\]

weakly as \(a_2 \to \infty\). Consequently,

\[
\frac{\mu - 2a_1}{2a_2} = \frac{\mu - 2a_1}{2\sqrt{a_1}} \frac{\sqrt{a_1}}{a_2} \to 0 \quad \text{and} \quad \frac{(\mu - 2a_1)^3}{a_2^3} = \left(\frac{\mu - 2a_1}{2\sqrt{a_1}}\right)^3 \frac{(2\sqrt{a_1})^3}{a_2^3} \to 0
\]
in probability as \(a_2 \to \infty\) and then \(\frac{(2a_1-3)^3}{a_2^3}h\left(\frac{2a_1}{a_2}\right)\) tends to 0 in probability as \(a_2 \to \infty\). For the term \(\frac{a_2-1}{8a_2^2}(\mu - 2a_1)^2\), similarly we have
\[
\frac{a_2-1}{8a_2^2}(\mu - 2a_1)^2 = \left(\frac{\mu - 2a_1}{2\sqrt{a_1}}\right)^2 \frac{a_1(a_2-1)}{2a_2^2} \to \frac{\sigma^2}{2} \chi^2_1
\]
weakly as \(a_2 \to \infty\). Putting all these limits into (3.14), we know
\[
\frac{f_\beta.a_1.a_2}{f_\beta.a_1} = e^{\log(K'_1L'_1(\mu))} \to \sqrt{1 + \sigma \exp\left\{-\frac{\sigma^2}{2} \chi^2_1\right\}}
\]
weakly as \(a_2 \to \infty\). By (3.12) and the Fatou Lemma,
\[
\lim_{a_2 \to \infty} \|\mathcal{L}(2a\lambda) - \mathcal{L}(\mu)\|_{TV} \geq \mathbb{E}\left|\sqrt{1 + \sigma e^{-\frac{\sigma^2}{2} \chi^2_1}} - 1\right| > 0.
\]
Finally the whole proof is closed now.

4 Appendix

Proof of Lemma 2.3. Review that \(x_i^2 \sim \chi^2_{(2a_1-\beta(i-1))}\) and \(y_i^2 \sim \chi^2_{\beta(m-(i-1))}\). Setting \(b_i := \beta m - 2\beta(i-1)\) for \(2 \leq i \leq m\) and \(b_1 = 0\), one gets
\[
z_i := x_i^2 + y_i^2 \sim \chi^2_{2a_1+b_i}
\]
for \(1 \leq i \leq m\) with the convention \(y_1 = 0\).

It is easy to see
\[
\sum_{i=1}^{m} b_i = \beta \sum_{i=1}^{m-1} (m - 2i) = 0;
\]
\[
\sum_{i=1}^{m} b_i^2 = \beta^2 \sum_{i=1}^{m-1} (m - 2i)^2 = \frac{\beta^2 m(m-1)(m-2)}{3}; \quad (4.1)
\]
\[
\sum_{i=1}^{m} b_i^3 = \beta^3 \sum_{i=1}^{m-1} (m - 2i)^3 = 0.
\]

Based on the Dumitriu and Edelman characterization, one gets
\[
\sum_{i=1}^{m} \mu_i = \text{tr}(AA') = \sum_{i=1}^{m} (x_i^2 + y_i^2) = \sum_{i=1}^{m} z_i. \quad (4.2)
\]

Therefore by Lemma 2.2 and (4.1), we have
\[
\mathbb{E} \sum_{i=1}^{m} \mu_i = \sum_{i=1}^{m} \mathbb{E}z_i = \sum_{i=1}^{m} (2a_1 + b_i) = 2a_1 m \quad (4.3)
\]
and consequently
\[
\operatorname{Var}\left(\sum_{i=1}^{m} \mu_i \right) = \sum_{i=1}^{m} \operatorname{Var}(z_i) = 2 \sum_{i=1}^{m} \mathbb{E}z_i = 4a_1m. \tag{4.4}
\]

By Dumitriu and Edelman’s characterization again, \(\mu_1 - 2a_1, \mu_2 - 2a_1, \ldots, \mu_m - 2a_1\) are the eigenvalues of the matrix \(AA' - 2a_1 I_m\). Therefore
\[
\sum_{i=1}^{m} (\mu_i - 2a_1)^2 = \operatorname{tr}((AA' - 2a_1 I_m)^2) = \sum_{i=1}^{m} (z_i - 2a_1)^2 + 2 \sum_{i=2}^{m} x_{i-1}^2 y_i^2. \tag{4.5}
\]

Remembering \(\mathbb{E}z_i = b_i + 2a_1\) for \(1 \leq i \leq m\), we have from Lemma 2.2
\[
\mathbb{E}(z_i - 2a_1)^2 = \mathbb{E}(z_i - \mathbb{E}z_i + b_i)^2 = \mathbb{Var}(z_i) + b_i^2 = 4a_1 + 2b_i + b_i^2
\]
for \(1 \leq i \leq m\). This, with (4.1), tells us
\[
\sum_{i=1}^{m} \mathbb{E}(z_i - 2a_1)^2 = 4a_1 m + \frac{\beta^2 m(m-1)(m-2)}{3}.
\]

By independence, it is clear that
\[
2 \sum_{i=2}^{m} \mathbb{E}(x_{i-1}^2 y_i^2) = 2\beta \sum_{i=2}^{m} (2a_1 - \beta(i-2))(m-(i-1))
= 2\beta a_1 m(m-1) - \frac{\beta^2 m(m-1)(m-2)}{3}.
\]

Then we have
\[
\mathbb{E} \sum_{i=1}^{m} (\mu_i - 2a_1)^2 = 4a_1 m(2 + \beta(m-1)) = 4a_1 mr.
\]

Now we investigate the third expression. It follows from (4.5) that
\[
\operatorname{Var}\left(\sum_{i=1}^{m} (\mu_i - 2a_1)^2 \right) = \sum_{i=1}^{m} \operatorname{Var}(z_i - 2a_1)^2 + 4 \sum_{i=2}^{m} \operatorname{Var}(x_{i-1}^2 y_i^2)
+ 4\operatorname{Cov}\left(\sum_{i=1}^{m} (z_i - 2a_1)^2, \sum_{i=2}^{m} x_{i-1}^2 y_i^2 \right). \tag{4.6}
\]

Next we examine one by one the three terms in (4.6). Using \(\mathbb{E}z_i = b_i + 2a_1\) and Lemma 2.2 again, we know
\[
\operatorname{Var}(z_i - 2a_1)^2 = \operatorname{Var}(z_i - \mathbb{E}z_i + b_i)^2
= \operatorname{Var}(z_i - \mathbb{E}z_i)^2 + 4b_i^2 \operatorname{Var}(z_i)
+ 4b_i \operatorname{Cov}(\mathbb{E}z_i)^2, z_i - \mathbb{E}z_i))
= 8\mathbb{E}z_i(\mathbb{E}z_i + 6) + 8b_i^2 \mathbb{E}z_i + 32b_i \mathbb{E}z_i
= 8(4a_1^2 + 12a_1 + (12a_1 + 6)b_i + (2a_1 + 5)b_i^2 + b_i^3).
By this expression and (4.1), we have

\[
\sum_{i=1}^{m} \text{Var}((z_i - 2a_1)^2) = 32a_1(a_1 + 3)m + \frac{8\beta^2}{3}(2a_1 + 5)m(m - 1)(m - 2). \tag{4.7}
\]

For the second term \(2 \sum_{i=1}^{m-1} x_i^2 y_i^{2} \) in (4.6), by independence, we have

\[
\begin{align*}
\text{Var}(x_i^2 y_i^{2}) &= \mathbb{E}[x_i^4 y_i^{4}] - (\mathbb{E}[x_i^2 y_i^{2}])^2 \\
&= ((\mathbb{E}x_i^2)^2 + 2\mathbb{E}x_i^2(\mathbb{E}y_{i+1}^2 + 2\mathbb{E}y_i^2)) - (\mathbb{E}x_i^2)^2(\mathbb{E}y_{i+1}^2)^2 \\
&= 2\mathbb{E}x_i^2\mathbb{E}y_i^2(\mathbb{E}x_i^2 + \mathbb{E}y_i^2 + 2) \\
&= 2(2a_1 + \beta - \beta i)(m - i)(2a_1 + \beta + 2 + \beta(m - 2i)). \tag{4.8}
\end{align*}
\]

Therefore it follows by careful calculation that

\[
\text{Var}\left(\sum_{i=1}^{m-1} x_i^2 y_i^{2}\right) = 4\beta a_1^2 m(m - 1) + 4\beta a_1 m(m - 1) + 2\beta^2 m(m - 1) \left( a_1 - \frac{m - 2}{3} \right). \tag{4.9}
\]

Now we work on the last term in (4.6). On the one hand, we have

\[
\begin{align*}
\text{Cov}((z_i - 2a_1)^2, y_i^2) &= \text{Cov}(y_i^2 - \mathbb{E}y_i^2 + x_i^2 - \mathbb{E}x_i^2 + b_i, y_i^2 - \mathbb{E}y_i^2) \\
&= \mathbb{E}(y_i^2 - \mathbb{E}y_i^2)^3 + 2b_i \text{Var}(y_i^2) \\
&= 4(2 + b_i)\mathbb{E}y_i^2 \tag{4.10}
\end{align*}
\]

for \(1 \leq i \leq m\). Here the last equality is guaranteed again by Lemma 2.2 and the second one is true since \(x_i^2 - \mathbb{E}x_i^2 + b_i\) is independent of \(y_i\). On the other hand, similarly we have

\[
\text{Cov}((z_i - 2a_1)^2, x_i^2) = 4(2 + b_i)\mathbb{E}x_i^2 \tag{4.11}
\]

for \(1 \leq i \leq m\). Therefore by independence and (4.10) and (4.11), we have

\[
\begin{align*}
\text{Cov}\left(\sum_{i=1}^{m} (z_i - 2a_1)^2, \sum_{i=2}^{m} x_i^2 y_i^{2}\right) &= \sum_{i=1}^{m-1} \mathbb{E}y_i^2 \text{Cov}((z_i - 2a_1)^2, x_i^2) + \sum_{i=2}^{m} \mathbb{E}x_i^2 \text{Cov}((z_i - 2a_1)^2, y_i^2) \\
&= 4 \sum_{i=1}^{m-1} \mathbb{E}y_i^2 \mathbb{E}x_i^2 (4 + b_i + b_{i+1}).
\end{align*}
\]
Thereby with simple algebra, we have

\[
\text{Cov}\left(\sum_{i=1}^{m}(z_i - 2a_1)^2, \sum_{i=2}^{m} x_{i-1}^2 y_i^2\right) = \frac{8}{3}(a_1 - 1)m(m - 1)(m - 2) + 16\beta a_1 m(m - 1).
\]

Plugging (4.7), (4.9) and (4.12) into (4.6), we finally have

\[
\text{Var}\left(\sum_{i=1}^{m}(\mu_i - 2a_1)^2\right) = 16\beta a_1^2 m(m - 1) + 16\beta^2 a_1 m(m - 1)(m - 2)
\]

\[
+ 8\beta^2 a_1 m(m - 1) + 80\beta a_1 m(m - 1)
\]

\[
+ 32a_1 m(a_1 + 3).
\]

Now we prove the expression for covariance. Similarly since all the random variables involved are independent, we have by (4.2) and (4.5),

\[
\text{Cov}\left(\sum_{i=1}^{m}(\mu_i - 2a_1)^2, \sum_{i=1}^{m}(\mu_i - 2a_1)\right) = \sum_{i=1}^{m} \text{Cov}\left((z_i - 2a_1)^2, x_i^2 + y_i^2\right)
\]

\[
+ 2\sum_{i=1}^{m-1} \text{Cov}(x_i y_{i+1}, x_i^2 + y_{i+1}^2).
\]

Then (4.10), (4.11), the independence of \(x_i\) and \(y_j\) and Lemma 2.2 show

\[
\text{Cov}\left(\sum_{i=1}^{m}(\mu_i - 2a_1)^2, \sum_{i=1}^{m}(\mu_i - 2a_1)\right) = 4\sum_{i=1}^{m} (b_i + 2a_1)(b_i + 2) + 8\sum_{i=1}^{m-1} \beta(m - i)(2a_1 - \beta(i - 1)).
\]

With simple calculus on the sum, we get from (4.1)

\[
\text{Cov}\left(\sum_{i=1}^{m}(\mu_i - 2a_1)^2, \sum_{i=1}^{m}(\mu_i - 2a_1)\right) = 8\beta a_1 m^2 + 8a_1 m(2 - \beta) = 16a_1 mr.
\]

It remains to prove the last expression. By the property of the random matrix \(A\), it is not hard to verify that

\[
\sum_{i=1}^{m}(\mu_i - 2a_1)^3 = \text{tr}\left((AA' - 2a_1 I_m)^3\right)
\]

\[
= \sum_{i=1}^{m}(z_i - 2a_1)^3 + 3 \sum_{i=1}^{m-1} x_i^2 y_{i+1}^2 (z_i + z_{i+1} - 4a_1).
\]

(4.13)
For the first term. It is ready to check that
\[
\mathbb{E}(z_i - \mathbb{E}z_i + b_i)^3 = \mathbb{E}(z_i - \mathbb{E}z_i)^3 + 3b_i\text{Var}(z_i) + b_i^3
\]
\[
= 8\mathbb{E}z_i + 6b_i\mathbb{E}z_i + b_i^3
\]
\[
= 8(2a_1 + b_i) + 6b_i(2a_1 + b_i) + b_i^3
\]
for \(1 \leq i \leq m\). Thereby with the help of (4.1), we have
\[
\mathbb{E} \sum_{i=1}^{m} (z_i - 2a_1)^3 = 16a_1m + 2\beta^2m(m - 1)(m - 2).
\]

Also by Lemma 2.2, it follows
\[
\mathbb{E}x_i^2y_{i+1}^2(z_i + z_{i+1} - 4a_1) = \mathbb{E}x_i^2y_{i+1}^2(x_i^2 + y_i^2 + x_{i+1}^2 + y_{i+1}^2 - 4a_1)
\]
\[
= \mathbb{E}x_i^2\mathbb{E}y_{i+1}^2(\mathbb{E}(x_i^2 + y_i^2 + 2) + \mathbb{E}(x_{i+1}^2 + y_{i+1}^2 + 2) - 4a_1)
\]
\[
= \mathbb{E}x_i^2\mathbb{E}y_{i+1}^2(b_i + b_{i+1} + 4)
\]
for any \(1 \leq i \leq m - 1\). Therefore by (4.12), we have
\[
3 \sum_{i=1}^{m-1} \mathbb{E}x_i^2y_{i+1}^2(z_i + z_{i+1} - 4a_1)
\]
\[
= 2\beta^2(a_1 - 1)m(m - 1)(m - 2) + 12\beta a_1m(m - 1).
\]

Consequently, we have
\[
\mathbb{E} \sum_{i=1}^{m} (\mu_i - 2a_1)^3 = 2\beta^2a_1m(m - 1)(m - 2) + 12\beta a_1m(m - 1) + 16a_1m.
\]

The proof is completed now. \(\square\)

**Proof of Lemma 2.4.** By (4.13) and the property of variance, we have
\[
\text{Var}\left(\sum_{i=1}^{m} (\mu_i - 2a_1)^3\right) \leq 2 \sum_{i=1}^{m} \text{Var}(z_i - 2a_1)^3
\]
\[
+ 18(m - 1) \sum_{i=1}^{m-1} \text{Var}(x_i^2y_{i+1}^2(z_i + z_{i+1} - 4a_1)).
\]
\[
(4.14)
\]

On the one hand, we have
\[
\text{Var}((z_i - 2a_1)^3) \leq \mathbb{E}(z_i - \mathbb{E}z_i + b_i)^6 \leq 32\mathbb{E}(z_i - \mathbb{E}z_i)^6 + 32b_i^6 = O(m^6),
\]
\[
(4.15)
\]
where the last equality holds since \( a_1 = O(m) \) and \( b_i = O(m) \) for all \( 1 \leq i \leq m \). On the other hand, it holds
\[
\text{Var}(x_i^2 y_{i+1}^2(z_i + z_{i+1} - 4a_1)) \leq 3\text{Var}(x_i^2 y_{i+1}^2(z_i - \mathbb{E}z_i)) \\
+ 3\text{Var}(x_i^2 y_{i+1}^2(z_{i+1} - \mathbb{E}z_{i+1})) \\
+ 3(b_i + b_{i+1})^2\text{Var}(x_i^2 y_{i+1}^2)
\]
for \( 1 \leq i \leq m - 1 \). By (4.8),
\[
3(b_i + b_{i+1})^2\text{Var}(x_i^2 y_{i+1}^2) = O(m^5).
\]
Obviously
\[
\text{Var}(x_i^2 y_{i+1}^2(z_i - \mathbb{E}z_i)) \leq \mathbb{E}x_i^4 y_{i+1}^4(z_i - \mathbb{E}z_i)^2 = \mathbb{E}y_{i+1}^4 \mathbb{E}x_i^4(z_i - \mathbb{E}z_i)^2.
\]
By the independence of \( x_i \) and \( y_i \), we have
\[
\mathbb{E}x_i^4(z_i - \mathbb{E}z_i)^2 = \mathbb{E}x_i^4(y_i^2 - \mathbb{E}y_i^2)^2 + \mathbb{E}x_i^4(x_i^2 - \mathbb{E}x_i^2)^2 \\
= \mathbb{E}x_i^4\text{Var}(y_i^2) + \mathbb{E}(x_i^2 - \mathbb{E}x_i^2)^4 \\
+ (\mathbb{E}x_i^2)^2\text{Var}(x_i^2) + 2\mathbb{E}x_i^2\mathbb{E}(x_i^2 - \mathbb{E}x_i^2)^3 \\
= 2\mathbb{E}x_i^4\mathbb{E}y_i^2 + 12\mathbb{E}x_i^4(\mathbb{E}x_i^2 + 4) + 2(\mathbb{E}x_i^2)^3 + 16(\mathbb{E}x_i^2)^2 \\
= O(m^3),
\]
where the third equality is guaranteed by Lemma 2.2. This ensures
\[
\text{Var}(x_i^2 y_{i+1}^2(z_i - \mathbb{E}z_i)) = O(m^5).
\]
Similarly we have
\[
\text{Var}(x_i^2 y_{i+1}^2(z_{i+1} - \mathbb{E}z_{i+1})) = O(m^5).
\]
Therefore
\[
\text{Var}(x_i^2 y_{i+1}^2(z_i + z_{i+1} - 4a_1)) = O(m^5). \quad (4.16)
\]
Hence putting (4.15) and (4.16) back into (4.14), we know
\[
\text{Var} \left( \sum_{i=1}^{m} (\mu_i - 2a_1)^3 \right) = O(m^7).
\]
The proof is completed. \( \square \)

**Acknowledgements**  This work was supported in part by the National Natural Science Foundation of China (Nos. 12171038, 11871008), the National Key R&D Program of China (No. 2020YFA0712900), 985 Projects the Fundamental Research Funds for the Central Universities (No. DUT20LK27).
References

1. Constatine A.G., Some non-central distribution problems in multivariate analysis. Ann. Math. Statist., 1963, 34: 1270–1285
2. Dumitriu I., Eigenvalues statistics for β-ensembles. Ph.D. Thesis, 1999
3. Dumitriu I., Edelman A., Matrix models for beta ensembles. J. Math. Phys., 2002, 43(11): 5830–5847
4. Dumitriu I., Edelman A., Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models. J. Math. Phys., 2006, 47: 063302, 36 pp.
5. Dumitriu I., Paquette E., Global fluctuations for linear statistics of β-Jacobi ensembles. Random Matrices: Theory Appl., 2012, 1(4): 1250013, 60 pp.
6. Edelman A., Sutton B., The β-Jacobi matrix model, the CS decomposition, and generalized singular value problems. Found. Comput. Math., 2008, 8(2): 259–285
7. Huang N., Ma Y., Limit theorems for β-Laguerre and β-Jacobi ensembles. Acta Math. Scientia (English Series), 2022, 42: 2025–2039
8. Jiang T., How many entries of a typical orthogonal matrix can be approximated by independent normals? Ann. Probab., 2006, 34(4): 1497–1529
9. Jiang T., Limit theorems for β-Jacobi ensembles. Bernoulli, 2013, 19(3): 1028–1046
10. Jiang T., Li D., Approximation of rectangular β-Laguerre ensembles and large deviations. J. Theor. Probab., 2015, 28: 804–847
11. Jiang T., Ma Y., Distances between random orthogonal matrices and independent normals. Trans. Amer. Math. Soc., 2019, 372: 1509–1553
12. Muirhead R.J., Aspects of Multivariate Statistical Theory. New York: Wiley, 1998