Limiting Empirical Spectral Distribution for Products of Rectangular Matrices

Yongcheng Qi[1], **Hongru Zhao**[2]

[1]Department of Mathematics and Statistics, University of Minnesota Duluth, 1117 University Drive, Duluth, MN 55812, USA. Email: yqi@d.umn.edu (corresponding author)

[2]Department of Mathematics and Statistics, University of Minnesota Duluth, 1117 University Drive, MN 55812, USA. Email: zhao1118@d.umn.edu.

Abstract. In this paper, we consider m independent random rectangular matrices whose entries are independent and identically distributed standard complex Gaussian random variables and assume the product of the m rectangular matrices is an n by n square matrix. We study the limiting empirical spectral distributions of the product where the dimension of the product matrix goes to infinity, and m may change with the dimension of the product matrix and diverge. We give a complete description for the limiting distribution of the empirical spectral distributions for the product matrix and illustrate some examples.

Keywords: Empirical spectral distribution, Eigenvalues, Product of rectangular matrices, Non-Hermitian random matrix

AMS 2010 Subject Classification: 15B52, 60F99, 60G57

1 Introduction

The study of Random Matrix Theory was initialized by Wishart [35] for statistical analysis of large samples. Wigner [34] found applications for random Hermitian matrix in nuclear physics. Subsequential applications include condensed matter physics (Beenakker [5]), number theory (Mezzadri and Snaith [22]), wireless communications (Couillet and Debbah [9], and high dimensional statistics (Johnstone [19, 20], Jiang [16]), quantum chromodynamics, chaotic quantum systems and growth processes (see, e.g., Akemann, Baik and Francesco [3]).

There are two major directions for the study for random matrices, including the empirical spectral distributions and the spectral radii. The classical semi-circular law was first introduced by Wigner, and then Ginibre [10] established the circle law for Ginibre ensembles. Since then, the assumptions were relaxed subsequently in the papers by Girko [11], Bai [4], Pan and Zhou [25], and Götze and Tikhomirov [14]. Tao and Vu [31] proved the circular law under the second moment condition. For the spectral radii, Tracy and Widom established the so-called Tracy-Widom laws for the limiting distributions for the three Hermitian matrices (Gaussian orthogonal ensemble, Gaussian unitary ensemble and Gaussian symplectic ensemble); see Tracy and Widom [32, 33]. Other work in this direct includes Rider [27, 28] and Rider and Sinclair [29].

Products of random matrices are particularly of interest in recent research. Ipsen [15] provided several applications, including wireless telecommunication, disordered spin chain, the stability of large complex system, quantum transport in disordered wires, symplectic maps and Hamiltonian mechanics, quantum chromo-dynamics at non-zero chemical potential. Götze and Tikhomirov [13], Bordenave [6], O'Rourke and Soshnikov [23] and O'Rourke *et al.* [24] found the limiting empirical spectral distribution for the product from the complex Ginibre ensemble when the number of multiplicands, say m, is fixed. Two recent papers by Jiang and Qi [17, 18] considered the spectral radii and limiting empirical spectral distribution for the product of complex Ginibre ensembles and the product of truncations of independent Haar unitary matrices by allowing m to change. Götze, Kösters and Tikhomirov [12], Zeng [36], and Chang and Qi [8] studied the limiting empirical distribution of product of the spherical ensemble. Chang, Li and Qi [7] investigated the limiting distribution of the spectral radii for product of matrices from the spherical ensemble.

In this paper, we consider the product of m random rectangular matrices with independent and identically distributed (i.i.d.) complex Gaussian entries and investigate the limiting empirical spectral distributions. Adhikari *et al.* [2] obtained the joint density function for the eigenvalues and found the limit of the expected empirical distributions when m is a fixed integer, and Zeng [37] obtained the limiting empirical spectral distribution. Lambert [21] established that the empirical distribution for square singular values converges to certain generalizations of the Fuss-Catalan distribution and that the maximum of the square singular values converges to the edge point of the Fuss-Catalan distribution. Very recently, Qi and Xie [26] obtained the limiting distributions for spectral radii for products of rectangular matrices when m changes with the dimension of the product matrices.

The rest of the paper is organized as follows. In Section 2, we introduce empirical spectral distributions for scaled eigenvalues from the production of independent random rectangular matrices and present a general result on the convergence of the empirical spectral distributions. We further investigate the limiting distributions and obtain all types of distributions and provide conditions when these distributions can be obtained. We also give a few illustrative examples. Proofs for the main results are given in Section 3.

2 Main Results

In this paper, we consider m independent rectangular matrices, $\mathbf{X}_j,\ 1 \le j \le m$, namely $\mathbf{X}_j$ is an $n_j \times n_{j+1}$ matrix for $1 \le j \le m$, where $n_1, \cdots, n_{m+1}$ are positive integers, and all entries of the m matrices are independent and identically distributed standard complex normal random variables. We assume $n_1 = n_{m+1} =: n$ so that the product

$$\mathbf{X}^{(m)} = \prod_{j=1}^{m} \mathbf{X}_j$$

is an $n \times n$ square matrix. We also assume $n = \min_{1\le j\le m+1} n_j$. In this case, the product matrix $\mathbf{X}^{(m)}$ is of full rank.

Denote the n eigenvalues of $\mathbf{X}$ as $\mathbf{z}_1, \cdots, \mathbf{z}_n$, and set $l_j = n_j - n \ge 0$, $j = 1, \cdots, m$. It follows from Theorem 2 of Adhikari *et al.* [2] that the joint density function for $\mathbf{z}_1, \cdots, \mathbf{z}_n$

is given by

$$p(z_1, \cdots, z_n) = C \prod_{1 \le j < k \le n} |z_j - z_k|^2 \prod_{j=1}^{n} w_m^{(l_1, \cdots, l_m)}(|z_j|) \tag{2.1}$$

with respect to Lebesgue measure on $\mathbb{C}^n$, where C is a normalizing constant such that $p(z_1, \cdots, z_n)$ is a probability density function, and function $w_m^{(l_1, \cdots, l_m)}(z)$ can be obtained recursively by

$$w_k^{(l_1, \cdots, l_k)}(z) = 2\pi \int_0^\infty w_{k-1}^{(l_1, \cdots, l_{k-1})}\left(\frac{z}{s}\right) w_1^{(l_k)}(s) \frac{ds}{s}, \quad k \ge 2$$

with initial $w_1^{(l)}(z) = \exp\left(-|z|^2\right) |z|^{2l}$ for any z in the complex plane; see Zeng [36].

Our objective in the paper is to investigate the limiting empirical spectral distribution of the product ensemble $\mathbf{X}^{(m)}$ when n tends to infinity. We allow m to change with n and substitute m_n for m from now on to show its dependence on n.

The empirical spectral distribution of $\mathbf{X}^{(m)}$ is the empirical distribution based on the eigenvalues, $\mathbf{z}_1, \cdots, \mathbf{z}_n$, of $\mathbf{X}^{(m)}$, i.e.,

$$\mu_n^* = \frac{1}{n} \sum_{j=1}^{n} \delta_{\mathbf{z}_j / a_n}, \tag{2.2}$$

where $a_n > 0$ is a sequence of normalizing constants. When m_n diverges with n, the magnitude of $\mathbf{z}_j$'s can go to infinity exponentially or vanish exponentially. In this case, one may not be able to find a sequence a_n such that the empirical measure μ_n^* converges. Instead, we will define empirical distribution for scaled eigenvalues as in Jiang and Qi [18].

Note that $\{\mathbf{z}_j;\, 1 \le j \le n\}$ are complex random variables. Write

$$\Theta_j = \arg(\mathbf{z}_j) \in [0, 2\pi) \;\text{ such that }\; \mathbf{z}_j = |\mathbf{z}_j| \cdot e^{i\Theta_j} \tag{2.3}$$

for $1 \le j \le n$. Further, assume that $Y_1, \cdots, Y_n$ are independent random variables and Y_j has a density function proportional to $y^{j-1} w_m^{(l_1, \cdots, l_m)}(y) I(y > 0)$. Given a sequence of positive measurable functions $h_n(r)$, $n \ge 1$, which are defined on $(0, \infty)$, we define the empirical measures for scaled eigenvalues as follows

$$\mu_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{(\Theta_j, h_n(|\mathbf{z}_j|))} \quad \text{and} \quad \nu_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{h_n(Y_j)}. \tag{2.4}$$

We note that the empirical spectral measure μ_n^* defined in (2.2) is the joint distribution for linearly scaled eigenvalues, which is the joint empirical distribution based on real parts and imaginary parts for linearly scaled eigenvalues. The empirical spectral measure μ_n defined in (2.4) is the joint distribution for arguments and scaled moduli of eigenvalues. The transformation h_n which applies to the moduli of eigenvalues can be any positive function. With notation in (2.3), we can use $(\Theta_j, h_n(\mathbf{z}_j))$ to form a new complex number $h_n(|\mathbf{z}_j|)e^{i\Theta_j}$. Therefore, we can define the empirical measure for scaled eigenvalues $h_n(|\mathbf{z}_j|)e^{i\Theta_j}$ as follows

$$\hat{\mu}_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{h_n(|\mathbf{z}_j|)e^{i\Theta_j}}. \tag{2.5}$$

We want to menton that two measures $\hat{\mu}_n$ and μ_n^* are the same when $h_n(r) = r/a_n$.

We will see later that the convergence of μ_n is equivalent to that of ν_n. In (2.4), if h_n is linear, that is, $h_n(r) = r/a_n$, where $\{a_n, n \geq 1\}$ is a sequence of positive numbers, we denote the empirical measure of $\mathbf{z}_j$'s by μ_n^* as in (2.2), and accordingly, we denote the empirical distribution of Y_j's by

$$\nu_n^* = \frac{1}{n} \sum_{j=1}^{n} \delta_{Y_j/a_n}. \tag{2.6}$$

We need the following notations as in the paper by Jiang and Qi [18].

• Any function $g(z)$ of complex variable $z = x + iy$, $x, y \in \mathbb{R}$ should be interpreted as a bivariate function of (x, y): $g(z) = g(x, y)$.

• We write $\int_A g(z)\, dz = \int_A g(x, y)\, dxdy$ for any measurable set $A \subset \mathbb{C}$.

• Unif(A) stands for the uniform distribution on a set A.

• For a sequence of random probability measures $\{\tau, \tau_n;\ n \geq 1\}$, we write

$$\tau_n \rightsquigarrow \tau \ \text{ if } \ \mathbb{P}(\tau_n \text{ converges weakly to } \tau \text{ as } n \to \infty)\text{=}1. \tag{2.7}$$

When τ is a non-random probability measure generated by random variable X, we simply write $\tau_n \rightsquigarrow X$. Review the notation "$\rightsquigarrow$" in (2.7). The symbol $\mu_1 \otimes \mu_2$ represents the product measure of two measures μ_1 and μ_2.

For determinantal point processes, Jiang and Qi [18] have established a general result on convergence of the empirical spectral distributions; see Lemma 3.1 in Section 3.

It follows form Lemma 3.1 that a common feature for limiting empirical distributions from determinant point processes is that the angle and radius of the random vector with the liming distribution are independent and the convergence of empirical distributions for the eigenvalues is equivalent to the convergence of the empirical distribution based the radii of the eigenvalues.

Inspired by Jiang and Qi [18] and Zeng [37], we define a sequence of distribution functions $F_n(x)$ as follows

$$F_n(x) = \Big(\prod_{j=1}^{m_n} \frac{nx+l_j}{n+l_j}\Big)^{1/\gamma_n} = \Big(\prod_{j=1}^{m_n}(1-\frac{n}{n_j}(1-x))\Big)^{1/\gamma_n}, \qquad x \in [0,1], \tag{2.8}$$

where $\{\gamma_n\}$ is a sequence of positive numbers to be selected so that F_n has a limit. Note that $F_n(x)$ is continuous and strictly increasing on $[0,1]$ with $F_n(0) = 0$ and $F_n(1) = 1$. It is easy to see that F_n is a distribution function on $[0,1]$. We assume $F_n(x) = 0$ when $x < 0$ and $F_n(x) = 1$ when $x > 1$.

We will assume that $F_n(x)$ converges weakly to a distribution function $F(x)$. This limit is closely related to the limiting empirical spectral distribution of μ_n and μ_n^* defined in (2.4) and (2.2).

A cumulative distribution F is a nondecreasing right-continuous function, and its generalized inverse defined as

$$F^*(u) = \inf\{x: \ F(x) > u\}, \quad u \in [0,1) \tag{2.9}$$

Define $F^*(u) = 0$ for $u < 0$ and $F^*(u) = 1$ for $u \geq 1$. One can show that $F^*(u)$ is also a nondecreasing right-continuous function and therefore, F^* is also a cumulative distribution function. When F is continuous and strictly increasing, F^* is the regular inverse of F.

Recall that F_n converges weakly to a distribution F if and only if $\lim_{n\to\infty} F_n(x) = F(x)$ for every continuity point x of F. A probability measure v is induced by F^* if $\nu((-\infty,u]) = F^*(u)$ for all u.

The main results of the paper are the following Theorems 2.1 and 2.2.

Theorem 2.1 *Let $\{m_n, n \geq 1\}$ be a sequence of positive integers and $\gamma_n > 0$. Assume that, for any positive integer k,*

$$c_k := \lim_{n\to\infty} \frac{1}{\gamma_n}\sum_{r=1}^{m_n}\left(\frac{n}{n_r}\right)^k \ \textit{exists} \tag{2.10}$$

with $c_1 \in (0, \infty)$, and $c_k \in [0, c_1]$ is non-increasing in $k \geq 2$. Define a distribution function F as follows

$$F(x) = \exp(-\sum_{k=1}^{\infty} \frac{c_k}{k}(1-x)^k), \qquad x \in (0,1], \tag{2.11}$$

and its generalized inverse, F^*, is given in (2.9). Set $h_n(x) = \frac{1}{a_n}|x|^{2/\gamma_n}$ with $a_n = \prod_{r=1}^{m_n} n_r^{1/\gamma_n}$. Then $\hat{\mu}_n \rightsquigarrow \mu$, where $\hat{\mu}_n$ is defined as in (2.5), and μ has a density function $\frac{f^*(|z|)}{2\pi|z|} I(F(0) \leq |z| \leq 1)$, where f^* is the density function of F^* and it can be also determined by $f^*(x) = 1/f(F^*(x))$ with $f(x) = F'(x)$, $x \in (0,1]$.

Theorem 2.2 *Let $\{m_n, n \geq 1\}$ be a sequence of positive integers and $\gamma_n > 0$. Assume*

$$\lim_{n\to\infty} \frac{1}{\gamma_n} \sum_{r=1}^{m_n} \frac{n}{n_r} = 0. \tag{2.12}$$

Define $h_n(x) = \frac{1}{a_n}|x|^{2/\gamma_n}$ with $a_n = \prod_{r=1}^{m_n} n_r^{1/\gamma_n}$. Then $\hat{\mu}_n \rightsquigarrow \mathrm{Unif}(|z| = 1)$, where $\hat{\mu}_n$ is defined as in (2.5).

Next, we present some general results on the convergence of the empirical distribution F_n. We will investigate the necessary and sufficient conditions for the weak convergence of F_n, characterize its limiting distribution F and reveal how the function F is related to the limit of the empirical measures μ_n. Theorems 2.1 and 2.2 are the direct consequences of the following two theorems.

Theorem 2.3 *Let $\{m_n\}$ be an arbitrary sequence of positive integers and $\{\gamma_n\}$ be a sequence of positive numbers such that F_n converges weakly to a probability distribution F. Let F^* denote the generalized inverse of F and ν be a probability measure induced by F^*. Define $a_n = \prod_{r=1}^{m_n} n_r^{1/\gamma_n}$ and $h_n(x) = \frac{1}{a_n}|x|^{2/\gamma_n}$ in (2.4). Then we have $\mu_n \rightsquigarrow \mathrm{Unif}[0, 2\pi) \otimes \nu$ as $n \to \infty$.*

Theorem 2.4 *Let $\{m_n\}$ be a sequence of positive integers, and γ_n be any sequence of positive numbers. If $F_n(x)$ converges weakly to a distribution function $F(x)$, then F is of one of the following three types*

(Type I). *$F(x)$ is continuous on $[0,1]$, and analytic on $(0,1)$, with $F(0+) \geq 0$, $F(1) = 1$, and the first derivative $f(x) = F'(x) > 0$ for $x \in (0,1)$;*

(Type II). *$F(0-) = 0$, $F(x) = 1$ for all $x \in [0, 1]$;*

(Type III). *$F(1) = 1$, $F(x) = 0$ for all $x \in [0, 1)$.*

Furthermore, we have

(a). $F_n(x)$ converges weakly to a Type I distribution if and only if condition (2.10) holds; Under condition (2.10), the limiting distribution F has a representation given in (2.11).

(b). $F_n(x)$ converges weakly to a Type II distribution if and only if (2.12) holds.

(c). $F_n(x)$ converges weakly to a Type III distribution if and only if

$$\lim_{n\to\infty} \frac{1}{\gamma_n} \sum_{r=1}^{m_n} \frac{n}{n_r} = \infty. \tag{2.13}$$

Remark 1 *From Theorem 2.4, we can draw the following conclusions.*
a. *If F is of type I, F is strictly increasing in $[0, 1]$ and its generalized inverse F^* is given by*

$$F^*(x) = \begin{cases} 0, & \text{if } x < F(0), \\ F^{-1}(x), & \text{if } x \in [F(0), 1), \\ 1, & \text{if } x \geq 1, \end{cases} \tag{2.14}$$

where the regular inverse F^{-1} of F is well defined over $[F(0), 1]$. F^* is continuous on $(-\infty, \infty)$ and strictly increasing on $[F(0), 1]$.
b. *If the limit F is of Type II, then its generalized inverse F^*, defined in (2.9), is given by*

$$F^*(x) = \begin{cases} 0, & \text{if } x < 1 \\ 1, & \text{if } x \geq 1. \end{cases} \tag{2.15}$$

This is a degenerate distribution at $x = 1$, that is, it induces a probability measure $\nu = \delta_1$, a delta function at 1. In this case, we have from Theorem 2.3 that $\mu_n \rightsquigarrow \mathrm{Unif}[0, 2\pi) \otimes \delta_1$. This is equivalent to that the empirical distribution $\hat{\mu}_n$ for scaled eigenvalues converges to the uniform distribution over the unit circle $|z| = 1$ in the complex plane; see Theorem 2.2.

c. *When F is of Type III, we have*

$$F^*(x) = \begin{cases} 0, & \textit{if } x < 0 \\ 1, & \textit{if } x \geq 0. \end{cases} \tag{2.16}$$

This defines a degenerate probability measure δ_0. Then the limit of $\hat{\mu}_n$ is degenerate at origin in the complex plane.

The most interesting case to us is the distribution of Type I; see Theorem 2.1. In this case, the normalization constant γ_n should be of the same order as $\sum_{j=1}^{m_n} \frac{n}{n_j}$, precisely, condition (2.10) must be true. One can simply take $\gamma_n = \sum_{j=1}^{m_n} \frac{n}{n_j}$ and calculate the limits when they exist

$$c_k := \lim_{n\to\infty} \frac{\sum_{j=1}^{m_n} \left(\frac{n}{n_j}\right)^k}{\sum_{j=1}^{m_n} \frac{n}{n_j}}, \qquad \text{for } k \geq 2.$$

Then we obtain the limiting distribution F via formula (2.11) with $c_1 = 1$. Type II and Type III limiting distributions can be trivially obtained by changing order of γ_n.

From Theorem 2.1, the limiting empirical distribution of $\hat{\mu}_n$ has a support on $F(0) \leq |z| \leq 1$. When $F(0) = 0$, the support is the unit disk. When $F(0) > 0$, $\{z : F(0) \leq |z| \leq 1\}$ is a ring. Since $F(x)$ is right continuous at $x = 0$, we have from (2.11) that $F(0) > 0$ if and only if $\sum_{k=1}^{\infty} \frac{c_k}{k} < \infty$. Some specific distributions on rings are given in Examples 2 and 3.

It is interesting to discuss when the empirical distribution μ_n^* for linearly scaled eigenvalues converges. This is equivalent to the convergence of μ_n or $\hat{\mu}_n$ when γ_n is set to be 2.

When m_n is actually a fixed integer, Zeng [37] obtained the liming distribution of μ_n^* by assuming that

$$\lim_{n\to\infty} \frac{n}{n_j} =: \alpha_j \in [0, 1], \qquad 2 \leq j \leq m; \tag{2.17}$$

see Theorem 1.1 in Zeng [37]. By selecting $\gamma_n = 2$, we can verify (2.10) holds, and

$$F(x) = x^{1/2} \prod_{j=2}^{m} \big(1 - \alpha_j(1 - x)\big)^{1/2}, \quad x \in (0, 1]. \tag{2.18}$$

Since $F(0) = 0$, the support of the liming distribution of μ_n^* is always the unit disk $\{z : |z| \leq 1\}$. With additional constraint $n = n_1 \leq n_2 \leq \cdots \leq n_m$, it is possible to show that (2.17) is also necessary for the convergence of F_n.

Consider the case $\lim_{n\to\infty} m_n = \infty$. By selecting $\gamma_n = 2$, (2.10) gives the necessary and sufficient conditions for convergence of μ_n^*. Again, in this case, $F(0) = 0$ for any limit F. If $\sum_{r=1}^{m_n} \frac{n}{n_j} \to \infty$, we can only consider the convergence of the empirical spectral distribution μ_n for nonlinearly scaled eigenvalues.

We offer one more comment as a remark before we give some illustrative examples.

Remark 2 *In Theorem 2.3, we have taken $h_n(r) = r^{2/\gamma_n}/a_n$ for $r > 0$ to re-scale the eigenvalues, where a_n is defined as $\prod_{j=1}^{m_n} n_j^{1/\gamma_n}$. As a matter of fact, if there exist some sequences $\gamma_n > 0$ and $a_n > 0$ such that $\mu_n \rightsquigarrow \mathrm{Unif}[0, 2\pi) \otimes \nu_1$ as $n \to \infty$, with $h_n(r) = r^{2/\gamma_n}/a_n$ and ν_1 being a non-degenerate probability measure, then we can show that*

$$\frac{\gamma_n}{\sum_{j=1}^{m_n} \frac{n}{n_j}} \to c \quad \textit{and} \quad \ln a_n - \frac{\sum_{j=1}^{m_n} \ln n_j}{\gamma_n} \to d \tag{2.19}$$

by using the laws of types, where $c \in (0, \infty)$ and $d \in (-\infty, \infty)$. This implies that there exist some sequences $\gamma_n > 0$ and $a_n > 0$ such that $\mu_n \rightsquigarrow \mathrm{Unif}[0, 2\pi) \otimes \nu_1$ with $h_n(r) = r^{2/\gamma_n}/a_n$ and ν_1 being a non-degenerate probability measure, if and only if $\mu_n \rightsquigarrow \mathrm{Unif}[0, 2\pi) \otimes \nu$, where ν is a non-degenerate probability measure with $h_n(r) = \left(r^2/\prod_{j=1}^{m_n} n_j\right)^{1/\sum_{j=1}^{m_n} n/n_j}$. Further, the relationship between ν and ν_1 under condition (2.19) is $\nu_1(-\infty, r] = \nu(-\infty, e^{cd} r^c]$ for all $r > 0$.

Example 1 *When these rectangular matrices are actually square matrices, that is, $n_1 = \cdots = n_{m_n+1} = n$, where m_n is any sequence of positive integers. Set $\gamma_n = m_n$. Then (2.10) holds trivially with $c_k = 1$ for all $k \geq 1$. We have $F(x) = x$, $x \in (0, 1]$. Then $G^*(x) = x$ for $x \in [0, 1]$ is the cumulative distribution function for uniform distribution over $[0, 1]$. This leads to Theorem 2 in Jiang and Qi [18].*

Example 2 *Let $\{m_n\}$ be positive integers such that $\lim_{n\to\infty} m_n = \infty$. Define $n_1 = n_{m_n+1} = n$ and assume $n_2 = \cdots = n_{m_n} \sim n\alpha_n$ as $n \to \infty$, where $\alpha_n \geq 1$. Then*

$$\lambda_k(n) := \sum_{j=1}^{m_n} \left(\frac{n}{n_j}\right)^k = 1 + \frac{m_n}{\alpha_n^k}(1 + o(1)) \qquad \textit{as } n \to \infty \tag{2.20}$$

for $k \geq 1$.

Assume $\lim_{n\to\infty} \alpha_n =: \alpha \in [1, \infty)$*. By taking* $\gamma_n = 2m_n$*, we see that* (2.10) *holds with* $c_1 = \frac{1}{2}$*, and* $c_1 = \frac{1}{2}\alpha^{-k}$ *for* $k \geq 1$*. We have*

$$F(x) = \Big(1 - \frac{1}{\alpha}(1-x)\Big)^{1/2} \qquad x \in (0, 1].$$

Then we obtain

$$F^*(x) = \begin{cases} 0, & \textit{if } x < \beta; \\ \frac{x^2-\beta^2}{1-\beta^2}, & \textit{if } x \in [\beta, 1); \\ 1, & \textit{if } x \geq 1 \end{cases}$$

with $\beta = (1 - \frac{1}{\alpha})^{1/2}$*. The density function of* F^* *is given by* $f^*(x) = \frac{2x}{1-\beta^2} I(\beta \leq x \leq 1)$*. According to Theorem 2.1,* $\hat{\mu}_n \rightsquigarrow \mathrm{Unif}(\beta \leq |z| \leq 1)$*. The limit is a uniform distribution on the ring* $\beta \leq |z| \leq 1$ *if* $\beta \in (0, 1)$*, and a uniform distribution on the unit disk if* $\beta = 0$*.*

Example 3 *In Example 2, we assume* $\lim_{n\to\infty} \alpha_n = \infty$*.*
(a)*. Consider the case* $\lim_{n\to\infty} \frac{m_n}{\alpha_n} = \infty$*. With selecting* $\gamma_n = \sum_{r=1}^{m_n} \frac{n}{n_j}$*, we have* $c_1 = 1$ *and* $c_k = 0$ *for all* $k \geq 2$*. Then we have* $F(x) = \exp(x - 1)$ *for* $x \in (0, 1]$*, yielding*

$$F^*(x) = \begin{cases} 0, & \textit{if } x < e^{-1}; \\ 1 + \ln x, & \textit{if } x \in [e^{-1}, 1); \\ 1, & \textit{if } x \geq 1. \end{cases}$$

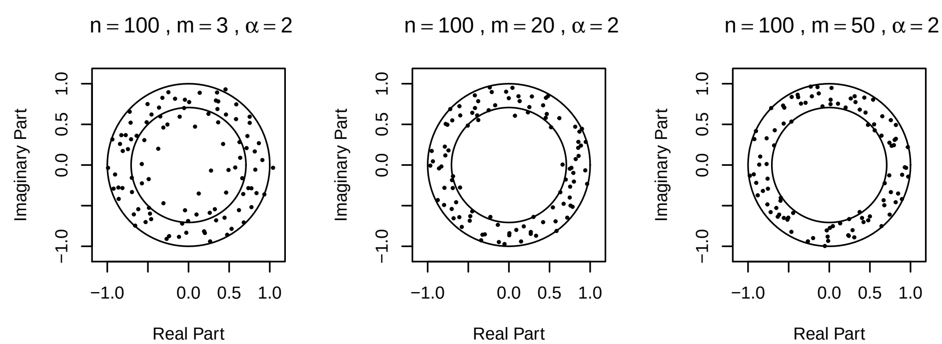
It follows from Theorem 2.1 that $\hat{\mu}_n \rightsquigarrow \mu$*, where* μ *has a density function* $\frac{1}{2\pi|z|^2} I(e^{-1} \leq |z| \leq 1)$*.*
(b)*. Consider the case* $\lim_{n\to\infty} \frac{m_n}{\alpha_n} = \gamma \in [0, \infty)$*. It follows from* (2.20) *that* $\lim_{n\to\infty} \lambda_1(n) = 1 + \gamma$*, and* $\lim_{n\to\infty} \lambda_k(n) = 1$ *for* $k \geq 2$*. This is the case we can establish the limiting law for* μ_n^**, the empirical distribution for linearly scaled eigenvalues, as defined in* (2.2)*. By selecting* $\gamma_n = 2$*, we have*

$$F(x) = x^{1/2} \exp\big(\frac{\gamma}{2}(x - 1)\big), \qquad x \in (0, 1].$$

Let f^* *denote the density of* $F^* = F^{-1}$ *on* $(0, 1))$*. We have* $\mu_n^* \rightsquigarrow \mu$*, where* μ *has a density function* $\frac{f^*(|z|)}{2\pi|z|}$*.*

To conclude this section, we carry out a simulation study by using the setup in Example 2. We select $\alpha_n = \alpha = 2$, $n_2 = \cdots = n_m = 2n$,and $\gamma_n = 2m$. Theoretically, if m is large, the empirical spectral distribution for the nonlinearly scaled eigenvalues is approximately uniformly distributed on the ring $\{\frac{\sqrt{2}}{2} \le |z| \le 1\}$. For each of $n = 100$ and $n = 400$, we select $m = 3$, $m = 20$ and $m = 50$ in order to see how well these scaled eigenvalues fit into the ring with the change in the value of m. The scatter plots for the scaled eigenvalues when $n = 100$ and $n = 400$ are given in Figures 1 and 2, respectively. From the two figures, we see that most of the scaled eigenvalues are already falling within the ring $\{\frac{\sqrt{2}}{2} \le |z| \le 1\}$ when $m = 20$.

Figure 1: Scatter plots for product matrices: $n_1 = n_{m+1} = n$, $n_2 = \cdots = n_m = 2n$, $\gamma_n = 2m$

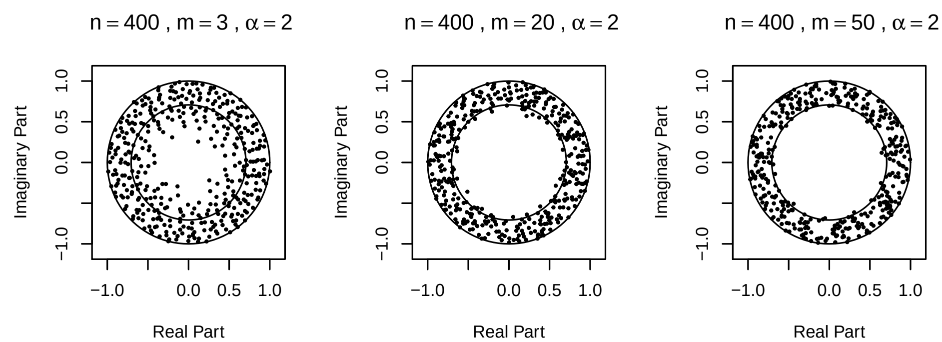


3 Proofs

The lemmas 3.1 and 3.2 below play a very important role in the proofs of our main results.

Lemma 3.1 *(Theorem 1 in Jiang and Qi [18]). Let $\varphi(x) \ge 0$ be a measurable function defined on $[0, \infty)$. Assume the density of $(Z_1, \cdots, Z_n) \in \mathbb{C}^n$ is proportional to $\prod_{1\le j<k\le n} |z_j - z_k|^2 \cdot \prod_{j=1}^{n} \varphi(|z_j|)$. Let $Y_1, \cdots, Y_n$ be independent r.v.'s such that the density of Y_j is proportional to $y^{2j-1}\varphi(y)I(y \ge 0)$ for every $1 \le j \le n$. Let μ_n,ν_n and ν_n^* be defined as in* (2.4) *and* (2.6)*, respectively. If $\{h_n\}$ are measurable functions such that $\nu_n \rightsquigarrow \nu$ for some probability*

Figure 2: Scatter plots for product matrices: $n_1 = n_{m+1} = n$, $n_2 = \cdots = n_m = 2n$, $\gamma_n = 2m$



measure ν, then $\mu_n \rightsquigarrow \mu$ with $\mu = \mathrm{Unif}[0, 2\pi] \otimes \nu$. Taking $h_n(r) = r/a_n$, the conclusion still holds if "(μ_n, ν_n, μ, ν)" is replaced by "$(\mu_n^*, \nu_n^*, \mu^*, \nu^*)$" where μ^* is the distribution of $Re^{i\Theta}$ with (Θ, R) having the law of $\mathrm{Unif}[0, 2\pi] \otimes \nu^*$.

Let $Y_1, \cdots, Y_n$ be the independent random variables determined in Lemma 3.1 under model (2.1). Let $\{s_{j,r},\ 1 \le j \le n,\ 1 \le r \le m_n\}$ be independent random variables and $s_{j,r}$ follow a Gamma$(l_r + j)$ with density function $y^{l_r+j-1}e^{-y}I_{y>0}/\Gamma\,(l_r + j)$. Set

$$T_j = \prod_{r=1}^{m_n} s_{j,r}, \quad 1 \le j \le n. \tag{3.1}$$

Lemma 3.2 *(Lemma 4 in Jiang and Qi [18]) Suppose $\{h_n(x);\, n \ge 1\}$ are measurable functions defined on $[0, \infty)$ and ν_n's are defined as in (2.4). Let $Y_1, \cdots, Y_n$ be as in Lemma 3.1 and ν be a probability measure on $\mathbb{R}$. Then $\nu_n \rightsquigarrow \nu$ if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mathbb{P}(h_n(Y_j) \le r) = H(r)$$

for every continuity point r of $H(r)$, where $H(r) := \nu((-\infty, r])$, $r \in \mathbb{R}$.

The results in the following lemma are summarized from Lemmas 2.2 and 2.3 from Zeng [37].

Lemma 3.3 *(Zeng [37]) We have*

$$\mathbb{P}(T_1 \le x) \ge \mathbb{P}(T_2 \le x) \ge \cdots \ge \mathbb{P}(T_n \le x) \tag{3.2}$$

for any $x \in [0, \infty)$, $(Y_1^2, \cdots, Y_n^2) \stackrel{d}{=} (T_1, \cdots, T_n)$, *and*

$$g(T_1, \cdots, T_n) \stackrel{d}{=} g(|z_1|^2, \cdots, |z_n|^2) \tag{3.3}$$

for any symmetric function $g(t_1, \cdots, t_n)$, *where* $\stackrel{d}{=}$ *denotes equality in distribution.*

Before we prove Theorem 2.4, we need to introduce more notation and preliminary results.

Define

$$\lambda_k(n) = \sum_{j=1}^{m_n} (\frac{n}{n_j})^k, \quad k \ge 1 \tag{3.4}$$

and

$$\theta_k(n) = \frac{\lambda_k(n)}{\lambda_1(n)}, \quad k \ge 1. \tag{3.5}$$

Note that $\lambda_k(n) \ge 1$ since $n_1 = n$. Since $n \le n_j$ for all $1 \le j \le m_n$ we have $\lambda_k(n)$ is non-increasing in $k \ge 1$ for each n. Thus, we have $\theta_k(n)$ is non-increasing in $k \ge 1$, implying that $0 < \theta_k(n) \le \theta_1(n) = 1$ for $k \ge 1$ and n.

We define a sequence of new distribution functions as follows

$$G_n(x) = \Big(\prod_{j=1}^{m_n} (1 - \frac{n}{n_j}(1-x)) \Big)^{1/\lambda_1(n)}, \quad x \in [0, 1] \tag{3.6}$$

These distributions are obtained by letting $\gamma_n = \lambda_1(n)$ in equation (2.8). We have

$$F_n(x) = G_n^{\lambda_1(n)/\gamma_n}(x), \quad x \in [0, 1]. \tag{3.7}$$

Set $g_n(x) = \ln(G_n(x))$, $x \in (0, 1]$. Then $g_n(x) \le 0$ for $x \in (0, 1]$. Using Taylor's expansion $\ln(1-t) = -\sum_{k=1}^{\infty} \frac{t^k}{k}$, $|t| < 1$, we have for $x \in (0, 1]$

$$g_n(x) = -\sum_{k=1}^{\infty} \frac{\theta_k(n)}{k}(1-x)^k. \tag{3.8}$$

Note that $\theta_1(n) = 1$ and $0 < \theta_k(n) \le 1$. We have the following inequalities

$$1 - x \le -g_n(x) \le \sum_{k=1}^{\infty} \frac{1}{k}(1-x)^k = -\ln(x), \quad x \in (0, 1]. \tag{3.9}$$

The probability distribution G_n over $[0,1]$ is defined via (3.6), and the same expression can not be extended beyond the interval $[0,1]$. The function $g_n(x)$, as the logarithm of G_n, has expansion (3.8) over $(0,1]$ only. However, g_n can be extended to a region in the complex plane via the expression on the right-hand side of (3.8). Now we fix $0<\delta<1$. For any complex number z such that $|z-1|\le\delta$, we have from (3.9) that

$$\sum_{k=1}^{\infty}\left|-\frac{\theta_k(n)}{k}(1-z)^k\right| \le \sum_{k=1}^{\infty}\frac{\theta_k(n)}{k}\delta^k = |g_n(1-\delta)| \le \ln(\frac{1}{1-\delta}). \tag{3.10}$$

Therefore, we can extent $g_n(x)$ to be a complex analytic function on disk $D=\{z\in\mathbb{C}: |z-1|<1\}$, namely

$$g_n(z) = -\sum_{k=1}^{\infty}\frac{\theta_k(n)}{k}(1-z)^k. \tag{3.11}$$

Lemma 3.4 *(Theorem 10.28 in Rudin [30]) Suppose f_j is analytic on open set $\Omega\subset\mathbb{C}$ for $j=1,2,\cdots$, and $f_j\to f$ uniformly on each compact subset of Ω. Then f is analytic on Ω, and $f_j'\to f'$ uniformly on any compact subset on Ω.*

Lemma 3.5 *Assume $\{n_s\}$ is a subsequence of $\{n\}$ such that $\lim_{s\to\infty}\theta_k(n_s)=a_k\in[0,1]$ for all $k\ge 2$. Set $a_1=1$. Then*

$$\lim_{s\to\infty} G_{n_s}(x) = G(x), \quad x\in(0,1], \tag{3.12}$$

where G is a distribution function given by

$$G(x) = \exp(-\sum_{k=1}^{\infty}\frac{a_k}{k}(1-x)^k), \quad x\in(0,1]. \tag{3.13}$$

$0<G(x)<1$ *is analytic and strictly increasing over* $(0,1)$.

Proof. In the proof we will use index n instead of n_s for the sake of brevity.

For each $\delta\in(0,1)$, set $K_\delta := \{z\in\mathbb{C}: |z-1|\le\delta\}$. It follows form (3.10) that $g_n(z)$ is uniformly bounded on K_δ.

Set $g(z) = -\sum_{k=1}^{\infty}\frac{a_k}{k}(1-z)^k$. The radius of convergence of $g(z)$ satisfies

$$R = \frac{1}{\limsup_{k\to\infty}\left(\frac{a_k}{k}\right)^{1/k}} \ge \frac{1}{\limsup_{k\to\infty}\left(\frac{1}{k}\right)^{1/k}} = 1,$$

i.e. $g(z)$ is well defined on disk $D = \{z \in \mathbb{C} : |z-1| < 1\}$. For each $\delta \in (0,1)$, we have

$$\sup_{z \in K_\delta} |g_n(z) - g(z)| \le \sum_{k=1}^{N} \frac{|a_k(n) - a_k|}{k} \delta^k + \sum_{k=N+1}^{\infty} \frac{\delta^k}{k} \le \sum_{k=1}^{N} |a_k(n) - a_k| + \frac{\delta^{N+1}}{1-\delta},$$

which implies

$$\begin{aligned}
\limsup_{n\to\infty} \sup_{z \in K_\delta} |g_n(z) - g(z)| &\le \limsup_{N\to\infty} \limsup_{n\to\infty} \Big(\sum_{k=1}^{N} |a_k(n) - a_k| + \frac{\delta^{N+1}}{1-\delta} \Big) \\
&\le \limsup_{N\to\infty} \frac{\delta^{N+1}}{1-\delta} \\
&= 0,
\end{aligned}$$

that is, $g_n(z)$ converges to $g(z)$ uniformly on K_δ.

Since $G_n(x) = \exp(g_n(x))$ for $x \in (0,1]$, we have (3.12) with $G(x) = \exp(g(x))$ for $x \in (0,1]$. Note that $g(x)$ is analytic, $g(x) < 0$ and is strictly increasing for $x \in (0,1)$, we have $0 < G(x) < 1$ is analytic and strictly increasing over $(0,1)$. ■

Lemma 3.6 *Assume $\{n_s\}$ is a subsequence of $\{n\}$ such that G_{n_s} converges weakly to a distribution G, then $\lim_{s\to\infty} \theta_k(n_s) =: a_k \in [0,1]$ for all $k \ge 2$, and G has a representation* (3.13)*.*

Proof. Note that $0 \le \theta_k(n) \le \theta_1(n) = 1$ for all $n \ge 1$ and $k \ge 2$. By the diagonal argument, for every subsequence of $\{n\}$, we can find its further subsequence along which $\theta_k(n)$ has a subsequential limit in $[0,1]$ for all $k \ge 2$.

We aim to show that $\lim_{s\to\infty} \theta_k(n_s)$ exists for all $k \ge 2$. If the conclusion is not true, then for some $k \ge 2$, say k_0, such that the limit of $\theta_{k_0}(n_s)$ doesn't exist. Then there exist two subsequences of $\{n_s\}$, say $\{n_{s'}\}$ and $\{n_{s''}\}$, such that

$$\lim_{s'\to\infty} \theta_{k_0}(n_{s'}) = a \ne b = \lim_{s''\to\infty} \theta_{k_0}(n_{s''}). \tag{3.14}$$

By the diagonal argument, we can find a further subsequence of $\{n_{s'}\}$, along which $\theta_k(n_{s'})$ has a subsequential limit $a_k \in [0,1]$ for each $k \ge 2$ with $a_{k_0} = a$. By Lemma 3.5 we have

$$G(x) = \exp(-\sum_{k=1}^{\infty} \frac{a_k}{k} (1-x)^k), \quad x \in (0,1]. \tag{3.15}$$

since any subsequential limit of $G_{n_s}(x)$ is equal to $G(x)$ in $x \in (0, 1]$. Similarly, we can find a further subsequence of $\{n_{s''}\}$, along which $\theta_k(n_{s''})$ has a subsequential limit $b_k \in [0, 1]$ for each $\geq k$ with $b_{k_0} = b$. Again, using Lemma 3.5 we have

$$G(x) = \exp(-\sum_{k=1}^{\infty} \frac{b_k}{k}(1-x)^k), \quad x \in (0, 1]. \tag{3.16}$$

By combining (3.15) and (3.16), we have

$$\sum_{k=1}^{\infty} \frac{a_k}{k}(1-x)^k = \sum_{k=1}^{\infty} \frac{b_k}{k}(1-x)^k, \quad x \in (0, 1].$$

Therefore, we $a_k = b_k$ for all $k \geq 2$, which contradicts $a_{k_0} = a \neq b = b_{k_0}$ from (3.14). This proves the lemma. ■

Proof of Theorem 2.4. First, we assume F_n converges weakly to a distribution function F. We will show F must be of one of the three types given in Theorem 2.4.

Review the definitions of $\lambda_k(n)$ and $\theta_k(n)$ in (3.4) and (3.5), respectively.

We consider the sequence $\lambda_1(n)/\gamma_n$. At this moment, we don't know yet whether $r_n := \lambda_1(n)/\gamma_n$ has a limit. We assume that $\{n_s\}$ is any subsequence of $\{n\}$ such that

$$\lim_{s\to\infty} r_{n_s} = \lim_{s\to\infty} \frac{\lambda_1(n_s)}{\gamma_{n_s}} = c_1 \in [0, \infty]. \tag{3.17}$$

We consider the following three cases individually: $c_1 \in (0, 1)$, $c_1 = \infty$, and $c_1 = 0$. From (3.7) we have that

$$G_n(x) = F_n^{1/r_n}(x) \qquad x \in (0, 1]. \tag{3.18}$$

Case 1. $c_1 \in (0, \infty)$ in (3.17).

In this case, we see that G_{n_s} converges weakly to $G(x) = F^{1/c_1}(x)$. By applying Lemma 3.6, we have $\lim_{s\to\infty} \theta_k(n_s) =: a_k \in [0, 1]$ for all $k \geq 2$, and G has a representation (3.13), which implies F has a representation (2.11) with $c_k = c_1 a_k$ for $k \geq 2$. This shows that F is of type I.

Case 2. $c_1 = 0$ in (3.17).

In view of (3.7), (3.8) and (3.9), we have for any $x \in (0, 1)$

$$1 \geq F_{n_s}(x) = G_{n_s}^{r_{n_s}}(x) = \exp(r_{n_s} g_{n_s}(x)) \geq \exp(r_{n_s} \ln(x)) \to 1$$

as $s \to \infty$, which implies $F(x) = 1$ for $x \in (0, 1)$ and thus F is of type II.

Case 3. $c_1 = \infty$ in (3.17).

Using the same equations as in the proof for *Case 2*, we have for any $x \in (0, 1)$

$$0 \le F_{n_s}(x) = G_{n_s}^{r_{n_s}}(x) = \exp(r_{n_s} g_{n_s}(x)) \le \exp(r_{n_s}(1 - x)) \to 0$$

as $s \to \infty$, which implies $F(x) = 0$ for $x \in (0, 1)$ and thus F is of type III.

We have proved that there are only three types of limiting distributions for F_n. Next, we will show the necessary and sufficient conditions in parts (a), (b), and (c).

Sufficiency for parts (b) and (c) has been proved. In fact, for part (b), condition (2.12) must be true when F is of Type II, otherwise, there exists a subsequential limit c_1 of r_n with $c_1 \in (0, \infty)$ or $c_1 = \infty$, such that F is of Type I or Type III, respectively, yielding a contradiction. A similar argument can be used to show (2.13) in part (c).

Finally, we need to prove part (a). The sufficiency has been proved in *Case 1* above. Assume F_n converges weakly to F, which is of Type I. We show (2.10), or equivalently, we show the following statements

Statement 1: $r_n = \lambda_1(n)/\gamma_n$ has a limit $c_1 \in (0, \infty)$;

Statement 2: For any $k \ge 2$, $\theta_k(n)$ has a finite limit.

If *Statement 1* is not true, then there are subsequences of $\{n\}$, say, $\{n_s\}$ and $\{n_{s'}\}$ such that

$$\lim_{s \to \infty} r_s = a \ne b = \lim_{s' \to \infty} r_{s'},$$

and $a, b \in (0, \infty)$. Any subsequential limit of r_n must be a finite positive number since F is of Type I.

From (3.18), we have G_{n_s} converges weakly to $F^{1/a}$ and $G_{n_{s'}}$ converges weakly to $F^{1/b}$. Then it follows from Lemma 3.6 that

$$F^{1/a}(x) = \exp(-\sum_{k=1}^{\infty} \frac{a_k}{k}(1 - x)^k), \quad x \in (0, 1]$$

and

$$F^{1/b}(x) = \exp(-\sum_{k=1}^{\infty} \frac{b_k}{k}(1 - x)^k), \quad x \in (0, 1]$$

where $a_1 = b_1 = 1$ and $a_k, b_k \in [0, 1]$ for all $k \ge 2$. We conclude that

$$a \sum_{k=1}^{\infty} \frac{a_k}{k}(1 - x)^k = b \sum_{k=1}^{\infty} \frac{b_k}{k}(1 - x)^k, \quad x \in (0, 1].$$

Since the functions on both sides of the above equation are analytic, their first derivatives at $x = 1$ must be the same, which leads to $a = b$, contradictory to the assumption $a \neq b$. Therefore, *Statement 1* is true, that is, r_n has a limit in $(0, \infty)$.

Given $\lim_{n\to\infty} r_n = c_1$, from (3.18) we have G_n converges weakly to F^{1/c_1}. Again, by using Lemma 3.6, we have $a_k := \lim_{n\to\infty} \theta_k(n)$ exists for all $k \geq 2$ and $a_k \in [0, 1]$. This proves *Statement 2*. The proof of the theorem is completed. ■

The following result is an extension of Lemma 2.3 in Zeng [37]. We allow m_n to change with n.

Lemma 3.7 *Assume $\{m_n\}$ is a sequence of positive integers. Then*

$$\frac{1}{\lambda_1(n)} \ln \frac{T_{[\mathrm{nx}]}}{\prod_{r=1}^{m_n}(l_r + n)} - \ln G_n(x) \xrightarrow{p} 0, \quad x \in (0, 1], \tag{3.19}$$

where $[nx]$ denotes the integer part of nx, T_j is defined in (3.1)*, $\lambda_1(n)$ is defined in* (3.4)*, and G_n is defined in* (3.6)*.*

Proof. We have $\ln T_j = \sum_{r=1}^{m} \ln s_{j,r}$ for $j \geq 1$. Since $s_{j,r}$ has a Gamma$(l_r + j)$ distribution, we have

$$\mu_{j,r} = \mathbb{E}\,(s_{j,r}) = l_r + j, \quad \mathrm{Var}\,(s_{j,r}) = l_r + j$$

and the moment generating function of $\ln s_{j,r}$ is

$$m_j(t) = \mathbb{E}\big(e^{t \ln s_{j,r}}\big) = \frac{\Gamma\,(l_r + j + t)}{\Gamma\big(l_r + j\big)}, \quad t > -(l_r + j).$$

It follows that

$$\mathbb{E}\big(\ln s_{j,r}\big) = \frac{d}{dt} m_j(t)\Big|_{t=0} = \frac{\Gamma'\,(l_r + j)}{\Gamma\,(l_r + j)} = \psi\,(l_r + j)\,,$$

where $\psi(x) = \Gamma'(x)/\Gamma(x)$ is a digamma function. Thus, we have

$$\mathbb{E}\big(\ln T_j\big) = \sum_{r=1}^{m_n} \mathbb{E}\,(\ln s_{j,r}) = \sum_{r=1}^{m_n} \psi\,(l_r + j)\,. \tag{3.20}$$

Set $\eta(t) = t - 1 - \ln t$ for $t > 0$. Then $\eta(t) \geq 0$ for $t > 0$.

Trivially, we have for any $1 \leq j \leq n$

$$\ln \frac{T_j}{\prod_{r=1}^{m_n} \mu_{j,r}} - \sum_{r=1}^{m_n} \ln\Big(\frac{l_r + j}{l_r + n}\Big) = \sum_{r=1}^{m_n}\Big(\frac{s_{j,r}}{\mu_{j,r}} - 1\Big) - \sum_{r=1}^{m_n} \eta\Big(\frac{s_{j,r}}{\mu_{j,r}}\Big). \tag{3.21}$$

It can be seen that for $1 \le j \le n$

$$\mathrm{Var}\left(\sum_{r=1}^{m_n}\Big(\frac{s_{j,r}}{\mu_{j,r}}-1\Big)\right)=\sum_{r=1}^{m_n}\frac{\mathrm{Var}\,(s_{j,r})}{\mu_{j,r}^2}=\sum_{r=1}^{m_n}\frac{1}{l_r+j}\le\sum_{r=1}^{m_n}\frac{n}{j(l_r+n)}=\frac{\lambda_1(n)}{j}.$$

Fix $x \in (0,1]$. Set $j = j_n = [nx]$. Then

$$\mathrm{Var}\Big(\frac{1}{\lambda_1(n)}\sum_{r=1}^{m_n}(\frac{s_{j_n,r}}{\mu_{j_n,r}}-1)\Big)\le\frac{1}{\lambda_1(n)j_n}\le\frac{1}{j_n}\to 0$$

as $n \to \infty$. By Chebyshev inequality, we obtain

$$\frac{1}{\lambda_1(n)}\sum_{r=1}^{m_n}(\frac{s_{j_n,r}}{\mu_{j_n,r}}-1)\xrightarrow{p}0. \tag{3.22}$$

From (3.21) and (3.20) we have

$$\mathbb{E}\Big(\sum_{r=1}^{m_n}\eta(\frac{s_{j,r}}{\mu_{j,r}})\Big)=\sum_{r=1}^{m_n}\ln\mu_{j,r}-\mathbb{E}\sum_{r=1}^{m_n}\ln s_{j,r}=-\sum_{r=1}^{m_n}\big(\psi(l_r+j)-\ln(l_r+j)\big).$$

We need the following approximation for ψ

$$\psi(t)-\ln t=-\frac{1}{2t}+O\Big(\frac{1}{t^2}\Big)\ \text{as}\ t\to\infty;$$

See, e.g., Formula 6.3.18 in Abramowitz and Stegun [1]. With $j_n = [nx]$, we have

$$\frac{1}{\lambda_1(n)}\mathbb{E}\Big(\sum_{r=1}^{m_n}\eta(\frac{s_{j,r}}{\mu_{j,r}})\Big)=\frac{O(1)}{\lambda_1(n)}\sum_{r=1}^{m_n}\frac{1}{l_r+j_n}\le\frac{O(1)}{\lambda_1(n)}\frac{\lambda_1(n)}{j_n}=O(\frac{1}{j_n})\to 0$$

as $n \to \infty$, which implies

$$\frac{1}{\lambda_1(n)}\sum_{r=1}^{m_n}\eta\left(\frac{s_{j,r}}{\mu_{j,r}}\right)\xrightarrow{p}0 \tag{3.23}$$

by Chebyshev inequality since $\sum_{r=1}^{m}\eta\left(\frac{s_{j,r}}{\mu_{j,r}}\right)\ge 0$. Therefore, combining (3.21), (3.22) and (3.23), we obtain

$$\frac{1}{\lambda_1(n)}\ln\frac{T_{[\mathrm{nx}]}}{\prod_{r=1}^{m_n}(l_r+[nx])}-\ln G_n(\frac{[nx]}{n})\xrightarrow{p}0,\quad x\in(0,1]. \tag{3.24}$$

From (3.8), we have $0 < g_n'(t) = \sum_{r=1}^{\infty}\theta_k(n)(1-t)^{k-1} \le \frac{1}{t}$ for $0 < t \le 1$, and hence,

$$|\ln G_n(\frac{[nx]}{n})-\ln G_n(x)|=|g_n(\frac{[nx]}{n})-g_n(x)|\le\sup_{\frac{[nx]}{n}\le t\le x}g_n'(t)|\frac{[nx]}{n}-x|\le\frac{1}{[nx]}\to 0$$

as $n \to \infty$. This, coupled with (3.24), yields (3.19). ■

Proof of Theorem 2.3. Assume F_n converges weakly to a distribution F. The conclusion in the theorem follows from Lemma 3.1 and Lemma 3.2 if we can prove

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\mathbb{P}(h_n(Y_j)\le y)=F^*(y)$$

for every continuity point y of F^*. According to (3.3), it is equivalent to show

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n}\le y\Big)=F^*(y) \tag{3.25}$$

for every continuity point y of F^*. Since $a_n=\prod_{r=1}^{m_n}n_r^{1/\gamma_n}=\prod_{r=1}^{m_n}(l_r+n)^{1/\gamma_n}$, we have

$$\mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n}\le y\Big)=\mathbb{P}\Big(\frac{1}{\gamma_n}\ln\frac{T_j}{\prod_{r=1}^{m_n}(l_r+j)}\le \ln y\Big),\quad y>0. \tag{3.26}$$

We also have the following two inequalities

$$\frac{j}{n}\mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n}\le y\Big)\le\frac{1}{n}\sum_{j=1}^{n}\mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n}\le y\Big)\le\frac{j}{n}+\mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n}\le y\Big), \tag{3.27}$$

which follow from the monotonicity in (3.2) directly.

<u>*Case 1*</u>. Assume F is of Type I.

In this case, F is strictly increasing in $[0,1]$ with $F(x)>0$ for any $x\in(0,1]$ and $\lim_{n\to\infty}F_n(x)=F(x)$, and F^* is given by (2.14). We note that $\lambda_1(n)/\gamma_n$ converges to a non-zero constant from Theorem 2.4. From (3.19), we have for any $x\in(0,1]$

$$\frac{1}{\gamma_n}\ln\frac{T_{[\mathrm{nx}]}}{\prod_{r=1}^{m_n}(l_r+n)}-\ln F_n(x)=\frac{\lambda_1(n)}{\gamma_n}\Big(\frac{1}{\lambda_1(n)}\ln\frac{T_{[\mathrm{nx}]}}{\prod_{r=1}^{m_n}(l_r+n)}-\ln G_n(x))\Big)\xrightarrow{p}0,$$

yielding

$$\frac{1}{\gamma_n}\ln\frac{T_{[\mathrm{nx}]}}{\prod_{r=1}^{m_n}(l_r+n)}-\ln F(x)\xrightarrow{p}0. \tag{3.28}$$

From (3.26) we get

$$\mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n}\le y\Big)=\mathbb{P}\Big(\frac{1}{\gamma_n}\ln\frac{T_j}{\prod_{r=1}^{m_n}(l_r+j)}-\ln F(x)\le\ln\frac{y}{F(x)}\Big) \tag{3.29}$$

for $y>0$ and any x with $F(x)>0$.

Now we are ready to show (3.25) when $F(0)<y<1$, $y\le F(0)$, and $y\ge 1$.

We first assume $F(0)<y<1$. Let $\delta\in(0,1)$ be any given number such that

$$F(0)<y-\delta<y<y+\delta<1.$$

Then $0 < F^*(y-\delta) < F^*(y) < F^*(y+\delta) < 1$.

By setting $x = F^*(y+\delta)$, $j = [nx]$ in (3.27) and (3.29) and using (3.28), we have

$$\begin{aligned}
& \limsup_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \le y\Big) \\
\le\ & \limsup_{n\to\infty} \frac{[nx]}{n} + \limsup_{n\to\infty} \mathbb{P}\Big(\frac{1}{\gamma_n}\ln\frac{T_{[nx]}}{\prod_{r=1}^{m_n}\Big(l_r+[nx]\Big)} - \ln F(x) \le \ln\frac{y}{y+\delta} < 0\Big) \\
=\ & x = F^*(y+\delta).
\end{aligned}$$

By setting $x = F^*(y-\delta)$ and $j = [nx]$ in (3.27) and (3.29), and using (3.28) again, we have

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \le y\Big) \ &\ge\ \liminf_{n\to\infty} \frac{[nx]}{n}\mathbb{P}\Big(\frac{1}{\gamma_n}\ln\frac{T_{[nx]}}{\prod_{r=1}^{m_n}(l_r+[nx])} - \ln F(x) \le \ln\frac{y}{y-\delta}\Big) \\
&=\ x = F^*(y-\delta).
\end{aligned}$$

Since $F^*(y)$ is continuous, we obtain (3.25) by letting δ tend to 0.

Assume $y \le F(0)$. For any $y_1 \in (F(0),1)$, we have

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \leqslant y\Big) \le \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \leqslant y_1\Big) = F^*(y_1),$$

which tends to $F^*(F(0)) = 0$ by letting $y_1 \downarrow F(0)$ since $F^*(y)$ is continuous. Similarly, when $y \ge 1$ we have for any $y_2 \in (F(0),1)$

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \leqslant y\Big) \ge \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \leqslant y_2\Big) = F^*(y_2),$$

which tends to $F^*(F(1)) = 1$ by letting $y_2 \uparrow 1$. In both cases, (3.25) still holds.

Case 2. Assume F is of Type II.

From Theorem 2.4, we have $\lim_{n\to\infty} \lambda_1(n)/\gamma_n = 0$. Since $\ln G_n(x) = g_n(x)$ is bounded for any fixed $x \in (0,1)$, we have from (3.19) that

$$\frac{1}{\gamma_n}\ln\frac{T_{[\mathrm{nx}]}}{\prod_{r=1}^{m_n}(l_r+n)} \xrightarrow{p} 0 \tag{3.30}$$

for any $x \in (0,1]$.

Review F^* in (2.15).

When $y \le 0$, (3.25) is trivially true.

When $y \in (0,1)$, we have $\ln y < 0$. For any $x \in (0,1)$, set $j = [nx]$. Then from (3.27), (3.26) and (3.30) we have

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \le y\Big) \le \limsup_{n\to\infty} \frac{[nx]}{n} + \limsup_{n\to\infty} \mathbb{P}\Big(\frac{1}{\gamma_n}\ln \frac{T_{[nx]}}{\prod_{r=1}^{m_n}\Big(l_r + [nx]\Big)} \le \ln y\Big) = x.$$

Since x can be arbitrarily small, we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \le y\Big) = 0 = F^*(y).$$

That is, (3.25) is true.

When $y > 1$, $\ln y > 0$. Again, for any $x \in (0,1)$, setting $j = [nx]$ and using (3.27), (3.26) and (3.30), we have

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \leqslant y\Big) \ge x.$$

By letting $x \uparrow 1$ we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{P}\Big(\frac{1}{a_n}T_j^{1/\gamma_n} \leqslant y\Big) = 1 = F^*(y).$$

This completes the proof of (3.25).

<u>*Case 3*</u>. Assume F is of Type III.

This time, we have $\lim_{n\to\infty} \lambda_1(n)/\gamma_n = \infty$.

$$\frac{1}{\gamma_n}\ln \frac{T_{[\mathrm{nx}]}}{\prod_{r=1}^{m_n}(l_r + n)} \xrightarrow{p} -\infty$$

for any $x \in (0,1)$. We can prove (3.25) by using similar lines to *Case 2*. We omit the details.

The proof of Theorem 2.3 is completed. ■

Proof of Theorem 2.1. From Theorems 2.3 and 2.4, we have $\mu_n \rightsquigarrow \mathrm{Unif}[0,2\pi) \otimes \nu$ as $n \to \infty$, where ν has a a density function $f^*(r) = \frac{d}{dr}F^*(r) = \frac{1}{f(F^*(r))}$, $r \in [F(0),1]$. Let Θ R are two independent random variables, Θ is uniformly distributed over $[0,2\pi)$ and R has density function f^*. Consider the transformation $Z = R\exp(i\Theta) = R\cos(\Theta) + iR\sin(\Theta) = (R\cos(\Theta), R\sin(\Theta)) =: (X,Y)$. Note that the Jacobian for transformation $(x,y) = (r\cos(\theta), r\sin(\theta))$ is $r = \sqrt{x^2+y^2} = |z|$, where $z = re^{i\theta} = x + iy$. The joint density function of $Z = (X, Y)$ is given by $\frac{1}{2\pi}\frac{f^*(|z|)}{|z|} = \frac{1}{2\pi f(F^*|z|)|z|} I(F(0) \le |z| \le 1)$. Since $\hat{\mu}_n$ is obtained under transformation $(x, y) = (r\cos(\theta), r\sin(\theta))$, by the continuous mapping theorem, we $\hat{\mu}_n$ converges with probability one to $Z = (X, Y)$ which has a joint density $\frac{1}{2\pi f(F^*|z|)|z|} I(F(0) \le |z| \le 1)$. ■

Proof of Theorem 2.2. Using the same notations as in the proof for Theorem 2.1, we have $P(R = 1) = 1$. Therefore, we can easily conclude that Z has a uniform distribution on the unit circle. ■

Acknowledgements. The authors would like to thank an anonymous referee for his/her careful reading of the manuscript and suggestion which has improved the layout of the manuscript. The research of Yongcheng Qi was supported in part by NSF Grant DMS-1916014.

References

[1] Abramowitz, M. and Stegun, I. A. (1972). Handbook of Mathematical Functions. *Dover, New York.*

[2] Adhikari, K., Reddy, N. K., Reddy, T. R. and Saha, K. (2016). Determinantal point processes in the plane from products of random matrices. *Ann. Inst. H. Poincare Probab. Statist.* 52 (1), 16-46.

[3] Akemann, G., Baik, J. and Francesco, P. D. (2011). The Oxford Handbook of Random Matrix Theory. *Oxford University Press, New York.*

[4] Bai, Z.D. (1997). Circular law. *Annals of Probability* 25, 494-529

[5] Beenakker, C. W. J. (1997). Random-matrix theory of quantum transport. *Rev. Mod. Phys.* 69, 731-809.

[6] Bordenave, C. (2011). On the spectrum of sum and product of non-Hermitian random matrices. *Elect. Comm. in Probab.* 16, 104-113.

[7] Chang, S., Li, D. and Qi, Y. (2018). Limiting distributions of spectral radii for product of matrices from the spherical ensemble. *Journal of Mathematical Analysis and Applications* 461, 1165-1176.

[8] Chang, S. and Qi, Y. (2017). Empirical distribution of scaled eigenvalues for product of matrices from the spherical ensemble. *Statistics & Probability Letters* 128, 8-13.

[9] Couillet, R. and Debbah, M. (2011). *Random matrix methods for wireless communications.* Cambridge Univ Press.

[10] Ginibre, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* 6, 440-449.

[11] Girko, V.L. (1984). The circular law. *Teoriya Veroyatnostei i ee Primeneniya* 29, 669-679.

[12] Götze F., Kösters H. and Tikhomirov, A. (2015). Asymptotic spectra of matrix-valued functions of independent random matrices and free probability. *Random Matrices: Theory Appl.* 4, 1550005.

[13] Götze, F. and Tikhomirov, T. (2010). On the asymptotic spectrum of products of independent random matrices. `http://arxiv.org/pdf/1012.2710v3.pdf`.

[14] Götze, F. and Tikhomirov, T. (2010). The circular law for random matrices. *Annals of Probability* 38, 1444-1491.

[15] Ipsen, J. R. (2015). Products of Independent Gaussian Random Matrices. Doctoral Dissertation. Bielefeld University.

[16] 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.

[17] Jiang, T. and Qi, Y. (2017). Spectral radii of large non-Hermitian random matrices. *Journal of Theoretical Probability* 30, 326-364.

[18] Jiang, T. and Qi, Y. (2019). Empirical distributions of eigenvalues of product ensembles. *Journal of Theoretical Probability* 32, 353-394,

[19] Johnstone, I. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Stat.* 29, 295-327.

[20] Johnstone, I. (2008). Multivariate analysis and Jacobi ensembles: Largest eigenvalue, Tracy-Widom limits and rates of convergence. *Ann. Stat.*, 36 (6), 2638-2716.

[21] Lambert, G. (2018). Limit theorem for biorthogonal ensembles and related combinatorial identities. *Adv. Math.* 329, 590-648.

[22] Mezzadri, F. and Snaith, N. C. (2005). *Recent perspectives in random matrix theory and number theory.* Cambridge Univ Press.

[23] O'Rourke, S. and Soshnikov, A. (2011). Products of independent non-Hermitian random matrices. *Electrical Journal of Probability* 16(81), 2219-2245.

[24] O'Rourke, S., Renfrew, D., Soshnikov, A. and Vu, V. (2015). Products of independent elliptic random matrices. *Journal of Statistical Physics* 160, 89-119.

[25] Pan, G. and Zhou, W. (2010). Circular law, extreme singular values and potential theory. *J. Multivariate Anal.* 101, 645-656.

[26] Qi, Y. and Xie, M. (2020). Spectral radii of products of random rectangular matrices. *Journal of Theoretical Probability* 33, 2185-2212.

[27] Rider, B. C. (2003). A limit theorem at the edge of a non-Hermitian random matrix ensemble. *J. Phys. A* 36(12), 3401-3409.

[28] Rider, B. C. (2004). Order statistics and Ginibre's ensembles. *Journal of Statistical Physics* 114, 1139-1148.

[29] Rider, B. C. and Sinclair, C. D. (2014). Extremal laws for the real Ginibre ensemble. *Ann. Appl. Probab.* 24(4), 1621-1651.

[30] Rudin, W. (1986). Real and Complex Analysis. 3rd edn. McGraw-Hill, New York.

[31] Tao, T. and Vu, V. (2010). Random matrices: universality of ESD and the circular law. *Annals of Probability* 38, 2023-2065.

[32] Tracy, C. A. and Widom, H. (1994). Level-spacing distributions and Airy kernal. *Comm. Math. Physics* 159, 151-174.

[33] Tracy, C. A. and Widom, H. (1996). On the orthogonal and symplectic matrix ensembles. *Comm. Math. Physics* 177, 727-754.

[34] Wigner, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann. Math.* 62, 548-564.

[35] Wishart, J. (1928). The generalized product moment distribution in samples from a normal multivariate population. *Biometrika* 20, 35-52.

[36] Zeng, X. (2016). Eigenvalues distribution for products of independent spherical ensembles. *J. Phys. A: Math. Theor.* 49, 235201.

[37] Zeng, X. (2017). Limiting empirical distribution for eigenvalues of products of random rectangular matrices. *Statistics and Probability Letters* 126, 33-40.