Triangular random matrices and biorthogonal ensembles

Dimitris Cheliotis [1]

August 12, 2016

Abstract

We study the singular values of certain triangular random matrices. When their elements are i.i.d. standard complex Gaussian random variables, the squares of the singular values form a biorthogonal ensemble, and with an appropriate change in the distribution of the diagonal elements, they give the biorthogonal Laguerre ensemble.

1 Introduction and statement of the results

1.1 Singular values of random matrices

Singular values of random matrices are of importance in numerical analysis, multivariate statistics, information theory, and the spectral theory of random non-symmetric matrices. See the survey paper Chafaï (2009). The starting point in this field is the result of Marchenko and Pastur (1967) (see also Theorem 3.6 in Bai and Silverstein (2010) for a more recent exposition), which is the following.

Let $\{X_{i,j} : i,j \in \mathbb{N}^+\}$ be i.i.d. complex valued random variables with variance 1, and for $n, m \in \mathbb{N}^+$ consider the $n \times m$ matrix $X(n,m) := (X_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m}$. Call $\lambda_1^{n,m} \geq \lambda_2^{n,m} \geq \cdots \geq \lambda_n^{n,m} \geq 0$ the eigenvalues of the Hermitian, positive definite matrix

$$S_{n,m} = \frac{1}{m} X(n,m) X(n,m)^*,$$

and

$$L_{n,m} := \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i^{n,m}}$$

their empirical distribution. Then for every $c > 0$, with probability 1, as $n, m \to \infty$ so that $n/m \to c$, $L_{n,m}$ converges weakly to the measure

$$\mathbf{1}_{a \leq x \leq b} \frac{1}{2\pi x c} \sqrt{(b-x)(x-a)} \, dx + \mathbf{1}_{c>1} \left(1 - \frac{1}{c}\right) \delta_0 \tag{1}$$

where $a = (1 - \sqrt{c})^2$, $b = (1 + \sqrt{c})^2$.

This is a universality result as the limit does not depend on the fine details of the distribution of the matrix elements $X_{i,j}$. On the other hand, the joint distribution of the eigenvalues $(\lambda_1^{n,m}, \lambda_2^{n,m}, \ldots, \lambda_n^{n,m})$, not surprisingly, depends on the exact distribution of the matrix elements. In a few cases this joint distribution can be determined. For example, if the $X_{i,j}$ follow the standard complex Gaussian distribution and $n \leq m$, the vector $(\lambda_1^{n,m}, \lambda_2^{n,m}, \ldots, \lambda_n^{n,m})$ has density with respect to Lebesgue measure in $\mathbb{R}^n$ which is

$$\frac{1}{\prod_{k=1}^n \Gamma(m-n+k)\Gamma(k)} e^{-\sum_{k=1}^n x_k} \left(\prod_{k=1}^n x_i\right)^{m-n} \prod_{1 \leq i \leq j \leq n} (x_i - x_j)^2 \mathbf{1}_{x_1 > x_2 > \cdots > x_n > 0}. \tag{2}$$

See, for example, relation (3.16) in Forrester (2010).

1.2   Triangular Gaussian matrices

In this work, we study the singular values of certain triangular random matrices. The motivation comes from the purely mathematical viewpoint as triangular matrices are ingredients in several matrix decompositions.

Assume as above that $\{X_{i,j} : i, j \in \mathbb{N}^+, i \geq j,\}$ are i.i.d. complex valued random variables with variance 1, and for $n \in \mathbb{N}^+$ let $X(n)$ be the lower triangular $n \times n$ matrix whose (i, j) element is $X_{i,j}$ for $1 \leq j \leq i \leq n$. Call $\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \cdots \geq \lambda_n^{(n)} \geq 0$ the eigenvalues of the Hermitian matrix

$$S_n = \frac{1}{n} X(n) X(n)^*,$$

and

$$L_n := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i^{(n)}}$$

their empirical distribution.

The fact that L_n converges weakly and description of the limit was given in Dykema and Haagerup (2004). It is analogous to the result of Marchenco and Pastur mentioned in the previous section and it says that with probability 1 the sequence $(L_n)_{n \geq 1}$ converges weakly to a deterministic measure μ_0 on $\mathbb{R}$ with moments

$$\int_{\mathbb{R}} x^k d\mu_0(x) = \frac{k^k}{(k+1)!} \tag{3}$$

for all $k \in \mathbb{N}$. The measure μ_0 is absolutely continuous with density that has support $[0, e]$ and can be expressed in terms of the Lambert function.

Here we do the obvious next step. That is, explore cases of distributions for the elements of the matrix $X(n)$ for which the joint distribution of the eigenvalues of $X(n)X(n)^*$ can be computed. The first such case is the following.

Theorem 1. *Let $n \in \mathbb{N}^+$ and assume that the random variables $\{X_{i,j} : i, j \in \mathbb{N}^+, i \geq j\}$ are complex standard normal. Then:*

(i). *The vector $\Lambda_n := (\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_n^{(n)})$ of the eigenvalues $\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \cdots \geq \lambda_n^{(n)}$ of $X(n)X^*(n)$ has density given by*

$$f_{\Lambda_n}(x_1, x_2, \ldots, x_n) = \frac{1}{\prod_{j=1}^{n-1} j!} e^{-\sum_{j=1}^n x_j} \prod_{i<j}(x_i - x_j)(\log x_i - \log x_j)\mathbf{1}_{x_1 > x_2 > \cdots > x_n > 0} \quad (4)$$

(ii). *The point process $\{\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_n^{(n)}\}$ is determinantal.*

The theorem will be implied by the more general Theorems 2 and 3 of the next subsection.

1.3 Eigenvalue realization of the biorthogonal Laguerre ensemble

The next model of random triangular matrix that we study is one where the elements strictly below the diagonal are i.i.d. standard complex normal but the elements of the diagonal are independent but not identically distributed.

More specifically, fix a positive integer n, reals $\theta \geq 0, b > 0$, and let

$$c_k = \theta(k-1) + b \quad (5)$$

for all $k \in \{1, 2, \ldots, n\}$. Next, consider the lower triangular matrix $X^{\theta,b}(n) = (X_{i,j})_{1 \leq i,j \leq n}$ with $\{X_{i,j} : 1 \leq j < i \leq n\}$ standard complex normal variables and $X_{k,k}$ having density

$$f_k(z) = \frac{1}{\pi \Gamma(c_k)} e^{-|z|^2} |z|^{2(c_k - 1)} \quad (6)$$

for all $z \in \mathbb{C}$. Thus $X_{k,k}$ can be written as

$$X_{k,k} = \frac{1}{\sqrt{2}} e^{i\phi_k} Y_k \quad (7)$$

where Y_k follows the χ_{2c_k} distribution and ϕ_k is uniform on $[0, 2\pi)$ independent of Y_k. For the distribution of the squares of the singular values of $X^{\theta,b}(n)$ we have the following theorem.

Theorem 2. *The vector $\Lambda_n := (\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_n^{(n)})$ of the eigenvalues $\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \cdots \geq \lambda_n^{(n)} \geq 0$ of $X^{\theta,b}(n)X^{\theta,b}(n)^*$ has density $f_{\Lambda_n}(x_1, x_2, \ldots, x_n)$ given by*

$$\frac{1}{\prod_{j=1}^{n-1} j!} \frac{\theta^{-n(n-1)/2}}{\prod_{k=1}^n \Gamma(c_k)} e^{-\sum_{j=1}^n x_j} \left(\prod_{j=1}^n x_j^{b-1} \right) \prod_{1 \leq i < j \leq n} (x_i - x_j)(x_i^\theta - x_j^\theta) \mathbf{1}_{x_1 > x_2 > \cdots > x_n > 0} \quad (8)$$

when $\theta > 0$, and

$$\frac{1}{\prod_{j=1}^{n-1} j!} \frac{1}{\Gamma(b)^n} e^{-\sum_{j=1}^{n} x_j} \left(\prod_{j=1}^{n} x_j^{b-1} \right) \prod_{1 \leq i < j \leq n} (x_i - x_j)(\log x_i - \log x_j) \mathbf{1}_{x_1 > x_2 > \cdots > x_n > 0} \qquad (9)$$

when $\theta = 0$.

Remark 1. i) When $\theta = 0$ and $b = 1$, the matrix $X^{\theta,b}(n)$ is exactly $X(n)$ of the previous subsection. And thus we get Part (i) of Theorem 1.

ii) When $\theta = 1$ and $b = m - n + 1$, with $m \geq n$ positive integers, (8) is the density (2). This is expected because there is a unitary matrix U so that $X(n,m)U \overset{d}{=} [X^{1,m-n}(n), 0]$, where 0 is the $n \times (m - n)$ zero matrix.

iii) The density in (8) is the density of the n-point biorthogonal Laguerre ensemble, so termed and studied in Section 4 of Borodin (1999), with parameter pair (α, θ) being $(b - 1, \theta)$. Note that (9) is the $\theta \to 0$ limit of (8).

iv) Densities of the form (8) were introduced by Muttalib (1995) in the context of disordered conductors. The conductance of such a conductor is given by the sum $\sum_{k=1}^{n}(1 + X_k)^{-1}$, where $\{X_k : 1 \leq k \leq n\}$ are random variables with joint density of the form

$$\prod_{1 \leq i < j \leq n} (x_i - x_j)(\{\operatorname{arcsinh}(\sqrt{x_i})\}^2 - \{\operatorname{arcsinh}(\sqrt{x_j})\}^2) \prod_{k=1}^{n} e^{-V_n(x_k)} \mathbf{1}_{x_k > 0}. \qquad (10)$$

See Beenakker and Rejaei (1993) for the derivation of this formula. With the purpose of having manageable formulas for the correlation functions of the X_i's, Muttalib suggested to simplify this density (10) by replacing the function $\{\operatorname{arcsinh}(\sqrt{x})\}^2$ by a polynomial $P_k(x)$. To illustrate the utility of this modification he considered the case of a monomial $P_k(x) = x^k$, he showed that the density then defines a determinantal point process, and gave formulas for the correlation functions. Later, Borodin (1999) gave an explicit formula for the kernel of the process for a few choices of the exponent V_n and using it determined the $n \to \infty$ limit of an appropriate blowup of the process around zero.

v) After the appearance of this preprint, Forrester et al. (2017) gave an alternative proof of Theorem 2. Moreover they show how utilizing the matrices $X^{\theta,b}(n)$ one can define a matrix whose eigenvalues have the density of the Jacobi biorthogonal ensemble, also defined in Borodin (1999).

The formula for f_{Λ_n} implies that $\{\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_n^{(n)}\}$ is a biorthogonal ensemble (Borodin (1999), Forrester (2010) Section 5.8). And this allows to prove with little effort that the ensemble is a determinantal point process. In the case $\theta > 0$, this is already known. We cover next the $\theta = 0$ case. Define

$$g_{j,k} := \int_0^\infty x^j (\log x)^k e^{-x} \, dx$$

for $j, k \in \mathbb{N}$, and consider the matrix $G := (g_{i,j})_{i,j \in \mathbb{N}}$.

Theorem 3. *For each positive integer n:*

(i). The matrix $G^{(n)} := (g_{j,k})_{0 \leq j,k \leq n-1}$ is invertible.

(ii). The point process $\{\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_n^{(n)}\}$ with law given by (9) is determinantal with kernel

$$K_n(x,y) = e^{-\frac{x+y}{2}}(xy)^{\frac{b-1}{2}} \sum_{j,k=1}^{n} c_{j-1,k-1}^{(n)} (\log y)^{j-1} x^{k-1}.$$

where $(c_{j,k}^{(n)})_{0 \leq j,k \leq n-1}$ is the inverse of $G^{(n)}$.

Finally, we come to the empirical spectral distribution $L_n^{\theta,b}$ of $X^{\theta,b}(n)X^{\theta,b}(n)^*/n$. The work in Dykema and Haagerup (2004) implies that this converges to a non trivial limit. To explain this connection, we need the notion of a DT-element.

Assume that ν a probability measure on $\mathbb{C}$ with compact support, and $c > 0$. For each n, let T_n be an $n \times n$ matrix with $(T_n)_{i,j} = 0$ if $1 \leq i \leq j \leq n$ and $\{(T_n)_{i,j} : 1 \leq j < i \leq n\}$ i.i.d. standard complex Gaussian. Also let D_n be a diagonal $n \times n$ matrix with i.i.d. diagonal elements, each having law ν, and independent of T_n. Finally, define $Z_n := D_n + cn^{-1/2}T_n$. It can be proved that for each $k \geq 1$ and $\varepsilon(1), \varepsilon(2), \ldots, \varepsilon(k) \in \{1, *\}$ the limit

$$\lim_{n \to \infty} \frac{1}{n} \mathbf{E}(\mathrm{tr}\{Z_n^{\varepsilon(1)} Z_n^{\varepsilon(2)} \cdots Z_n^{\varepsilon(k)}\}) \tag{11}$$

exists (Theorem 2.1 in Dykema and Haagerup (2004)).

Definition 1. An element x of a $*$-noncommutative probability space $(\mathcal{A}, \phi)$ is called a $DT(\nu, c)$-element if for every $k \geq 1$ and $\varepsilon(1), \varepsilon(2), \ldots, \varepsilon(k) \in \{1, *\}$, we have that $\phi(x^{\varepsilon(1)} x^{\varepsilon(2)} \cdots x^{\varepsilon(k)})$ equals the value in (11).

And we are now ready to discuss the convergence of the sequence $(L_n^{\theta,b})_{n \geq 1}$.

Theorem 4. *The empirical distribution of the eigenvalues of $X^{\theta,b}(n)X^{\theta,b}(n)^*/n$ converges to a measure μ_θ whose moments are the moments of xx^* where x is a $DT(\nu_\theta, 1)$ element, and ν_θ is the uniform measure on the disc $D(0, \sqrt{\theta}) := \{z \in \mathbb{C} : |z| \leq \sqrt{\theta}\}$.*

Note that the limit does not depend on b. In the case that $\theta > 1$ and $b = 1$, it is proven in Paragraph 4.5.1 of Claeys and Romano (2014) that the measure μ_θ has density f_θ with support $I_\theta = [0, (1+\theta)^{1+1/\theta}]$. To describe it, let $J : \mathbb{C}\backslash[-1,0] \to \mathbb{C}$ with

$$J(z) = (z+1)\left(\frac{z+1}{z}\right)^{1/\theta} \theta.$$

For each x interior point of I_θ, there are exactly two solutions of $J(z) = x$, which are conjugate complex numbers. Call them $I_-(x), I_+(x)$ so that $\mathrm{Im}(I_+(x)) > 0$. Then the density f_θ is given by

$$f_\theta(x) = \begin{cases} \frac{\theta}{2\pi xi}(I_+(x) - I_-(x)) & \text{if } x \in \left(0, (1+\theta)^{1+1/\theta}\right), \\ 0 & \text{if } x \in \mathbb{R}\backslash\left(0, (1+\theta)^{1+1/\theta}\right). \end{cases} \tag{12}$$

Orientation: Theorems 2, 3, 4 are proved in Sections 2, 3, 4 respectively.

2 Distribution of singular values for $X^{\theta,b}(n)$

Define the following sets of matrices

$\mathcal{T}_n$: lower triangular $n \times n$ matrices with elements in $\mathbb{C}$ and diagonal elements in $\mathbb{C} \backslash \{0\}$.

$\mathcal{T}_n^+$: elements of $\mathcal{T}_n$ with diagonal elements in $(0, \infty)$,

$\mathcal{V}_n$: diagonal $n \times n$ matrices with diagonal elements complex of modulus 1.

$\mathcal{M}_n^+$: positive definite $n \times n$ matrices with elements in $\mathbb{C}$.

We identify the spaces $\mathcal{T}_n, \mathcal{T}_n^+, \mathcal{V}_n$ with $\mathbb{R}^{n(n-1)} \times (\mathbb{R}^2 \backslash \{0,0\})^n, \mathbb{R}^{n(n-1)} \times (0, \infty)^n, [0, 2\pi)^n$ respectively, and view $\mathcal{M}_n^+$ as a subset of $n \times n$ Hermitian matrices, which we identify with $\mathbb{R}^{n^2}$. Densities of random variables with values in these spaces are meant with respect to the corresponding Lebesgue measure.

Consider the maps $g : \mathcal{T}_n^+ \times \mathcal{V}_n \to \mathcal{T}_n, h : \mathcal{T}_n^+ \to \mathcal{M}_n^+$ with

$$g(T,V) := TV, \tag{13}$$

$$h(Y) := YY^*. \tag{14}$$

They are both one to one and onto. Call $g^{-1} := (\gamma_1, \gamma_2)$, and $X := X^{\theta,b}(n)$. Then $XX^* = h(\gamma_1(X))$ provided that $X \in \mathcal{T}_n$, which holds with probability 1. We will use this relation in order to find the joint law of the elements of XX^*, and then, the law of its eigenvalues will follow from a well known formula.

Lemma 1. *The Jacobian of the map g has absolute value*

$$\prod_{j=1}^n t_{j,j}.$$

Proof. Let $X := g(T,V) = TV$ and call $x_{i,j}$ its (i,j) element. For a complex number x, we write x^R, x^I for its real and imaginary part respectively. The Jacobian matrix of g is an $n(n+1) \times n(n+1)$ block diagonal matrix with n blocks, one for each column of X. I.e., it is of the form

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & A_n \end{pmatrix}. \tag{15}$$

The block A_j, corresponding to column j, is the $\{2(n-j+1)\} \times \{2(n-j+1)\}$ matrix

$$\frac{\partial(x_{j,j}^R, x_{j,j}^I, x_{j+1,j}^R, x_{j+1,j}^I, \ldots, x_{n,j}^R, x_{n,j}^I)}{\partial(\theta_j, t_{j,j}, t_{j+1,j}^R, t_{j+1,j}^I, \ldots, t_{n,j}^R, t_{n,j}^I)},$$

and an easy computation shows that its determinant equals $-t_{j,j}$. ∎

Lemma 2. *The map h has Jacobian*

$$2^n \prod_{i=1}^n t_{i,i}^{2(n-i)+1}.$$

Proof. This is Proposition 3.2.6 of Forrester (2010). ∎

In the following, we use the notation set in Subsection 1.3. Let $C(\theta, b) := \left(\prod_{k=1}^{n} \Gamma(c_k)\right)^{-1}$. The density of $X^{\theta,b}(n)$ is

$$f_{X^{\theta,b}(n)}(x) = \frac{1}{\pi^{n(n+1)/2}} C(\theta, b)\, e^{-\sum_{1 \le j \le i \le n} |x_{i,j}|^2} \prod_{k=1}^{n} |x_{k,k}|^{2(c_k-1)} \tag{16}$$

$$= \frac{1}{\pi^{n(n+1)/2}} C(\theta, b)\, e^{-\operatorname{tr}(xx^*)} \prod_{k=1}^{n} |x_{k,k}|^{2(c_k-1)} \tag{17}$$

for all $x \in \mathbb{C}^{n(n+1)/2}$.

For an $n \times n$ matrix $a = (a_{i,j})_{1 \le i,j \le n}$ and $k \in \{1, 2, \ldots, n\}$, we denote by a_k its main $k \times k$ minor, that is, the matrix $(a_{i,j})_{1 \le i,j \le k}$.

Proposition 1. *Let $X := X^{\theta,b}(n)$. The matrix $A := XX^*$ has density*

$$f_A(a) = \frac{1}{\pi^{n(n-1)/2}} \frac{1}{\prod_{k=1}^{n} \Gamma(c_k)}\, e^{-\operatorname{tr}(a)} \{\det(a)\}^{c_n - 1} \{\det(a_1)\det(a_2)\cdots\det(a_{n-1})\}^{-\theta - 1} \tag{18}$$

for all $a \in \mathcal{M}_n^+$, and $f_A(a) = 0$ for every Hermitian matrix not an element of $\mathcal{M}_n^+$.

Proof. Let $(T, V) := g^{-1}(X)$. Since $XX^* = h(T)$, our first step is to find the distribution of T. The density of the pair (T, V) is

$$f_{T,V}(t, v) = f_X(g(t,v))|Jg(t,v)| = f_X(t) \prod_{j=1}^{n} t_{j,j}$$

for $(t, v) \in \mathcal{T}_n^+ \times \mathcal{V}_n$. We used that $f_X(tv) = f_X(t)$ for all $v \in \mathcal{V}_n$. We integrate $f_{T,V}(t, v)$ over v to find the marginal of T as

$$f_T(t) = (2\pi)^n f_X(t) \prod_{j=1}^{n} t_{j,j}.$$

Now, for given $a \in \mathcal{M}_n^+$, let $t := h^{-1}(a)$. Then

$$f_A(a) = f_T(h^{-1}(a))|Jh^{-1}(a)| = (2\pi)^n f_X(h^{-1}(a)) \left(\prod_{j=1}^{n} t_{j,j}\right) \frac{1}{|Jh(h^{-1}(a))|}$$

$$= (2\pi)^n \frac{1}{\pi^{n(n+1)/2}} C(\theta, b)\, e^{-\operatorname{tr}(a)} \prod_{j=1}^{n} |t_{j,j}|^{2(c_j - 1)} \frac{1}{2^n \prod_{j=1}^{n} t_{j,j}^{2(n-j)+1}} \prod_{j=1}^{n} t_{j,j}$$

$$= \frac{1}{\pi^{n(n-1)/2}} C(\theta, b)\, e^{-\operatorname{tr}(a)} \left(\prod_{i=1}^{n} t_{j,j}^{2(n-j-c_j+1)}\right)^{-1}$$

$$= \frac{1}{\pi^{n(n-1)/2}} C(\theta, b)\, e^{-\operatorname{tr}(a)} \left(\prod_{j=1}^{n} t_{j,j}^2\right)^{c_n - 1} \left(\prod_{j=1}^{n} t_{j,j}^{2(n-j)}\right)^{-(1+\theta)}.$$

In the third equality we used Lemma 2, and in the last equality the fact that $-c_j = \theta(n-j) - c_n$ for all $j \in \{1, 2, \ldots, n\}$. Finally, we express the products involving the variables $t_{j,j}$ in terms of the variable a. Since T is lower triangular, we have $a_i = T_i T_i^*$. Thus

$$\det(a_i) = |\det(T_i)|^2 = (t_{1,1}t_{2,2}\ldots t_{i,i})^2.$$

Multiplying these equalities for all $1 \le i \le n-1$, we get

$$\det(a_1)\det(a_2)\cdots\det(a_{n-1}) = \prod_{i=1}^{n} t_{i,i}^{2(n-i)}.$$

This finishes the proof of the proposition. ∎

Proof of Theorem 2. From relations (4.1.17), (4.1.18) in Anderson et al. (2010), and the fact that $X^{\theta,b}(n)X^{\theta,b}(n)^*$ is positive definite, we have that the vector of the eigenvalues in decreasing order has density

$$f_{\Lambda_n}(\lambda) = C_n \prod_{i<j}(\lambda_i - \lambda_j)^2 \int_{U(n)} f_A(HD_\lambda H^*)(dH)\, \mathbf{1}_{\lambda_1 > \lambda_2 > \ldots > \lambda_n > 0}$$

where $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_n)$, D_λ is the diagonal matrix with diagonal λ, (dH) is the normalized Haar measure on $U(n)$, and the constant C_n is

$$C_n := \frac{\pi^{n(n-1)/2}}{\prod_{j=1}^{n-1} j!}.$$

Thus, writing $a := HD_\lambda H^*$ and taking into account Proposition 1, we get

$$f_{\Lambda_n}(\lambda) = \frac{C(\theta, b)}{\prod_{j=1}^{n-1} j!}\left\{\prod_{i<j}(\lambda_i - \lambda_j)^2\right\} e^{-\sum_{j=1}^{n}\lambda_j}\left(\prod_{j=1}^{n}\lambda_j\right)^{c_n-1} K(\lambda)\, \mathbf{1}_{\lambda_1 > \lambda_2 > \ldots > \lambda_n > 0} \qquad (19)$$

with

$$K(\lambda) := \int_{U(n)} \{\det(a_1)\det(a_2)\cdots\det(a_{n-1})\}^{-\theta-1}\,(dH). \qquad (20)$$

The computation of the last integral is given in Lemma 3. Combining that computation with (19), we finish the proof. ∎

Lemma 3. *For $\theta \ge 0$, the integral in (20) equals*

$$K(\lambda) = \frac{\prod_{1 \le i < j \le n}\int_{\lambda_j}^{\lambda_i} x^{-\theta-1}\,dx}{\prod_{1 \le i < j \le n}(\lambda_i - \lambda_j)} = \left(\prod_{i=1}^{n}\lambda_i\right)^{-\theta(n-1)}\frac{\prod_{1 \le i < j \le n}\int_{\lambda_j}^{\lambda_i} x^{\theta-1}\,dx}{\prod_{1 \le i < j \le n}(\lambda_i - \lambda_j)}. \qquad (21)$$

Proof. To simplify the exposition, we introduce a binary relation which we denote by $\succ$. $x \succ y$ means that there is $k \in \mathbb{N}^+$ so that $x = (x_1, x_2, \ldots, x_{k+1}) \in \mathbb{R}^{k+1}, y = (y_1, y_2, \ldots, y_k) \in \mathbb{R}^k$, and

$$x_1 \ge y_1 \ge x_2 \ge y_2 \cdots \ge x_n \ge y_n \ge x_{k+1}.$$

8

For $x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$ with $x_1 \geq x_2 \geq \ldots \geq x_k$, we let $Y(x)$ be the set of all elements $(y^{(k-1)}, y^{(k-2)}, \ldots, y^{(1)})$ of $\mathbb{R}^{k-1} \times \mathbb{R}^{k-2} \times \cdots \times \mathbb{R}^2 \times \mathbb{R}$ that satisfy

$$x \succ y^{(k-1)} \succ y^{(k-2)} \succ \cdots \succ y^{(2)} \succ y^{(1)}.$$

It is shown in Lemma 1.12 of Baryshnikov (2001) that

$$\text{Vol}(Y(x)) = \prod_{1 \leq i < j \leq k} \frac{x_i - x_j}{j - i}. \tag{22}$$

We can now start the proof of the lemma. For each $i = 1, 2, \ldots, n-1$, call $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, \cdots, x_i^{(i)})$ the vector of the eigenvalues of the symmetric matrix a_i with $x_1^{(i)} \geq x_2^{(i)} \geq \cdots \geq x_i^{(i)}$. Then the integrand in (20) is simply

$$\prod_{i=1}^{n-1} \prod_{j=1}^{i} (x_j^{(i)})^{-\theta-1}.$$

Under (dH), the law of a is the one of an $n \times n$ GUE matrix conditioned to have eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, and according to Proposition 4.7 in Baryshnikov (2001), the law of $(x^{(n-1)}, x^{(n-2)}, \ldots, x^{(1)})$ is the uniform on $Y(\lambda)$ with respect to Lebesgue measure. Thus the integral equals

$$\frac{1}{\text{Vol}(Y(\lambda))} \underbrace{\int_{\lambda_n}^{\lambda_{n-1}} \int_{\lambda_{n-1}}^{\lambda_{n-2}} \cdots \int_{\lambda_2}^{\lambda_1}}_{n-1 \text{ integrals}} \underbrace{\int_{x_{n-1}^{(n-1)}}^{x_{n-2}^{(n-1)}} \cdots \int_{x_2^{(n-1)}}^{x_1^{(n-1)}}}_{n-2 \text{ integrals}} \cdots \underbrace{\int_{x_2^{(2)}}^{x_1^{(2)}}}_{1 \text{ integral}} \prod_{1 \leq j \leq i \leq n-1} (x_j^{(i)})^{-\theta-1} \prod_{1 \leq j \leq i \leq n-1} dx_j^{(i)}. \tag{23}$$

Assume now that $\theta > 0$. Let $\tilde{\lambda}(\theta) := (\lambda_n^{-\theta}, \lambda_{n-1}^{-\theta}, \ldots, \lambda_1^{-\theta})$, and in the integral make the change of variables $y_j^{(i)} = (x_j^{(i)})^{-\theta}$ for all $1 \leq j \leq i \leq n-1$. Then the previous expression becomes

$$\frac{1}{\text{Vol}(Y(\lambda))} (-\theta)^{-n(n-1)/2} \underbrace{\int_{\lambda_n^{-\theta}}^{\lambda_{n-1}^{-\theta}} \int_{\lambda_{n-1}^{-\theta}}^{\lambda_{n-2}^{-\theta}} \cdots \int_{\lambda_2^{-\theta}}^{\lambda_1^{-\theta}}}_{n-1 \text{ integrals}} \underbrace{\int_{y_{n-1}^{(n-1)}}^{y_{n-2}^{(n-1)}} \cdots \int_{y_{n-1}^{(n-1)}}^{y_1^{(n-1)}}}_{n-2 \text{ integrals}} \cdots \underbrace{\int_{y_2^{(2)}}^{y_1^{(2)}}}_{1 \text{ integral}} \prod_{i=1}^{n-1} \prod_{j=1}^{i} dy_j^{(i)}$$

$$= \theta^{-n(n-1)/2} \frac{\text{Vol}(Y(\tilde{\lambda}(\theta)))}{\text{Vol}(Y(\lambda))} = \theta^{-n(n-1)/2} \frac{\prod_{1 \leq i < j \leq n} (\lambda_j^{-\theta} - \lambda_i^{-\theta})}{\prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)} \tag{24}$$

$$= \left(\prod_{i=1}^{n} \lambda_i \right)^{-\theta(n-1)} \frac{\theta^{-n(n-1)/2} \prod_{1 \leq i < j \leq n} (\lambda_i^{\theta} - \lambda_j^{\theta})}{\prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)}. \tag{25}$$

And the lemma is proved in this case. In the case $\theta = 0$, in the integral of (23), we let $y_j^{(i)} = \log x_j^{(i)}$ for all $1 \leq j \leq i \leq n-1$ and proceed as above. Alternatively, we can take $\theta \to 0$ in the last expression. ∎

3 Determinantal process. Proof of Theorem 3

Proof. (i). For each positive integer n and $y_1, y_2, \ldots, y_n \in \mathbb{R}$, we let

$$\Delta(y_1, y_2, \ldots, y_n) := \det(y_k^{j-1})_{1 \leq j, k \leq n} = \prod_{1 \leq j < k \leq n} (y_k - y_j).$$

9

Equation (3.3) of Deift and Gioev (2009) gives that the determinant of $G^{(n)}$ is

$$\det\left(\int_0^\infty x^j(\log x)^k e^{-x}\,dx\right)_{0\le j,k<n} = \frac{1}{n!}\int_0^\infty\cdots\int_0^\infty \det(x_k^{j-1})_{1\le j,k\le n}\det((\log x_k)^{j-1})_{1\le j,k\le n}\prod_{k=1}^n dx_k \tag{26}$$

$$= \frac{1}{n!}\int_0^\infty\cdots\int_0^\infty \Delta_n(x_1,\ldots,x_n)\Delta_n(\log x_1,\ldots,\log x_n)\prod_{k=1}^n dx_k \tag{27}$$

The integrand is positive, thus the determinant is not zero.

(ii) Follows from part (i) and Proposition 5.8.1 of Forrester (2010). ∎

In the rest of the section, we discuss the structure of $G^{(n)}$ and compute explicitly the value of its determinant.

Lemma 4. *The matrix $G := (g_{i,j})_{i,j\in\mathbb{N}}$ has an LU factorization $G = LU$ with*

$$L_{i,j} = |s(i+1,j+1)| \qquad\qquad \text{for } i\ge j\ge 0, \tag{28}$$

$$U_{i,j} = (j)_i g_{0,j-i} \qquad\qquad \text{for } 0\le i\le j. \tag{29}$$

Here $s(\cdot,\cdot)$ denotes the Stirling number of the first kind while $(j)_i$ is the falling factorial defined as $(x)_i := x(x-1)\cdots(x-(i-1))$ for all $x\in\mathbb{C}$ and i positive integer, and $(x)_0 = 1$ for all $x\in\mathbb{C}$.

For the the definition of the Stirling numbers, see Chapter 13 in van Lint and Wilson (2001). Since $L_{k,k} = 1$ and $U_{k,k} = k!$ for all $k\in\mathbb{N}$, we obtain

$$\det(G^{(n)}) = 1!\,2!\cdots(n-1)!.$$

Proof. We compute the exponential generating function of the sequence $(g_{j,k})_{j,k\in\mathbb{N}}$.

$$\sum_{j,k=0}^\infty \frac{u^j}{j!}\frac{v^k}{k!}g_{j,k} = \int_0^\infty e^{-x}e^{ux}e^{v\log x}\,dx = \int_0^\infty e^{-x}e^{-(1-u)x}x^v\,dx = (1-u)^{-v-1}\Gamma(1+v)$$

$$= \sum_{j=0}^\infty (v+1)(v+2)\cdots(v+j)\frac{u^j}{j!}\sum_{s=0}^\infty \frac{\Gamma^{(s)}(1)}{s!}v^s$$

$$= \sum_{j=0}^\infty \frac{u^j}{j!}\sum_{r=0}^j (-1)^{j-r}s(j+1,r+1)v^r\sum_{s=0}^\infty \frac{\Gamma^{(s)}(1)}{s!}v^s$$

$$= \sum_{j,k=0}^\infty \frac{u^j}{j!}v^k\sum_{r=0}^{j\wedge k}\frac{\Gamma^{(k-r)}(1)}{(k-r)!}(-1)^{j-r}s(j+1,r+1).$$

Since for $m,n\in\mathbb{N}$ the integer $s(m,n)$, if not zero, has sign $(-1)^{m-n}$ and $\Gamma^{(n)}(1) = g_{0,n}$, we get

$$g_{j,k} = \sum_{r=0}^{j\wedge k}|s(j+1,r+1)|(k)_r g_{0,k-r}$$

for all $j,k\in\mathbb{N}$. This proves the factorization $G = LU$. ∎

10

4 Limiting empirical distribution in the $\theta > 0$ case

Proof of Theorem 4: We first prove the following.

CLAIM: The sequence $\left(X^{\theta,b}(n)/\sqrt{n}\right)_{n\geq 1}$ converges in $*$-moments to a $DT(\nu_\theta, 1)$ element.
Recall the form of the diagonal elements of $X^{\theta,b}(n)$ in (7). Denote by

$\mu_{n,k}$: the law of $X_{k,k}/\sqrt{n}$,

$\eta_n := \mu_{n,1} \times \mu_{n,2} \times \cdots \times \mu_{n,n}$, the law of the vector $(X_{k,k}/\sqrt{n})_{1\leq k\leq n}$,

$\tilde\eta_n$: the symmetrization of η_n, that is, the law of $(X_{\pi(k),\pi(k)}/\sqrt{n})_{1\leq k\leq n}$, where π is a random permutation of $\{1, 2, \ldots, n\}$ uniformly chosen and independent of the matrix,

$\tilde\eta_n^{(p)}$: the marginal of the first p coordinates of $\tilde\eta_n$ ($p \in \{1, 2, \ldots, n\}$).

According to Theorem 2.13 in Dykema and Haagerup (2004), it is enough to show that $\tilde\eta_n^{(p)}$ converges in $*$-moments to $\times_1^p \nu_\theta$, where $\times_1^p \nu_\theta$ is the product measure having p factors each equal to ν_θ. We prove this in three steps.

STEP 1. $\lim_{n\to\infty} d_{TV}(\tilde\eta_n^{(p)}, \times_1^p \tilde\eta_n^{(1)}) = 0$.

Here, d_{TV} is the total variation distance for probability measures. Since

$$\tilde\eta_n^{(p)} = \frac{1}{(n)_p} \sum_{\substack{r_1,r_2,\ldots,r_p\in[n] \\ \text{all different}}} \mu_{n,r_1} \times \mu_{n,r_2} \times \cdots \times \mu_{n,r_p} \leq \frac{n^p}{(n)_p} \times_1^p \tilde\eta_n^{(1)},$$

(recall that $(n)_p$ is the falling factorial, defined in the statement of Lemma 4) the function

$$q_n := \frac{n^p}{(n)_p} \times_1^p \tilde\eta_n^{(1)} - \tilde\eta_n^{(p)}$$

is a measure and

$$\tilde\eta_n^{(p)} - \times_1^p \tilde\eta_n^{(1)} = \left(\frac{n^p}{(n)_p} - 1\right) \times_1^p \tilde\eta_n^{(1)} - q_n.$$

The right hand side of the last equality is the difference of two measures each having mass $n^p/(n)_p - 1$. Thus the total variation distance between $\tilde\eta_n^{(p)}$ and $\times_1^p \tilde\eta_n^{(1)}$ is at most $2(n^p/(n)_p - 1)$ which tends to 0 as $n \to \infty$.

STEP 2. $\tilde\eta_n^{(1)}$ converges weakly to ν_θ as $n \to \infty$.

Note that if X has distribution ν_θ, then $|X|^2$ is uniformly distributed on $[0, \theta]$. Since the law of $\tilde\eta_n^{(1)}$ is radially symmetric, it suffices to prove that for any bounded Lipschitz function $h : \mathbb{R} \to \mathbb{R}$, it holds that

$$\frac{1}{n} \sum_{k=1}^n \mathbf{E}h\left(\frac{Y_k^2}{2n}\right) \to \frac{1}{\theta} \int_0^\theta h(r)\, dr \tag{30}$$

as $n \to \infty$. Take independent random variables $(W_k)_{k\geq 1}$ so that $W_1 \sim \Gamma(b, 1)$ and $W_k \sim \Gamma(\theta, 1)$ for all $k \geq 2$. Then $Y_k^2/2$ has the same law as $S_k := W_1 + W_2 + \cdots + W_k$. The left hand side of (30) is

$$\frac{1}{n} \sum_{k=1}^n \mathbf{E}h\left(\frac{S_k}{k}\frac{k}{n}\right). \tag{31}$$

Call I_n this quantity and $\hat{I}_n := n^{-1} \sum_{k=1}^{n} h(\theta k/n)$, which converges to $\int_0^1 h(\theta x)\,dx$ since h is continuous. Then $|I_n - \hat{I}_n| \leq J_n := n^{-1} \sum_{k=1}^{n} \mathbf{E}(\min\{2||h||_\infty, C|S_k/k - \theta|\})$ where C is a Lipschitz constant for h. Using the fact that S_k/k converges pointwise to θ and the bounded convergence theorem, we get $\lim_{k\to\infty} \mathbf{E}(\min\{2||h||_\infty, C|S_k/k - \theta|\}) = 0$. Thus, $\lim_{n\to\infty} J_n = 0$ and (30) is proved.

STEP 3. $\tilde{\eta}_n^{(p)}$ converges in $*$-moments to $\times_1^p \nu_\theta$ as $n \to \infty$.

It follows from the previous two steps and the Cramer-Wold theorem that $\tilde{\eta}_n^{(p)}$ converges weakly to $\times_1^p \nu_\theta$. This implies convergence in $*$-moments because for all $k_1, k_2, \ldots, k_p \in \mathbb{N}^+$ it holds

$$\sup_{\substack{r_1, r_2, \ldots, r_p \in [n] \\ \text{all different}}} \int |x_1|^{2k_1} |x_2|^{2k_2} \cdots |x_p|^{2k_p} \, d\mu_{n,r_1}(x_1)\, d\mu_{n,r_2}(x_2) \cdots d\mu_{n,r_p}(x_k) < \infty.$$

The last assertion is easy to show using the formulas for the moments of the chi squared distribution.

Now call x the $\mathrm{DT}(\nu_\theta, 1)$ element mentioned in the claim. We can assume that x is an element of a von Neumann algebra (see Remark 2.3 in Dykema and Haagerup (2004)), and thus there is a unique measure μ_θ which has support a compact subset of $\mathbb{R}$ and moments the same as xx^* (Lemma 5.2.19 in Anderson et al. (2010)). The theorem follows by combining this with the above claim. ∎

References

G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20 of *Springer Series in Statistics*. Springer, 2010.

Y. Baryshnikov (2001). GUEs and queues. *Probab. Theory Related Fields*, 119(2):256–274.

C. Beenakker and B. Rejaei (1993). Nonlogarithmic repulsion of transmission eigenvalues in a disordered wire. *Physical Review Letters*, 71(22):3689–3692.

A. Borodin (1999). Biorthogonal ensembles. *Nuclear Phys. B*, 536(3):704–732.

D. Chafaï (2009). Singular values of random matrices. *Lecture Notes*. Available at author's website.

T. Claeys and S. Romano (2014). Biorthogonal ensembles with two-particle interactions. *Non-linearity*, 27(10):2419.

P. Deift and D. Gioev. *Random matrix theory: invariant ensembles and universality*, volume 18 of *Courant Lecture Notes in Mathematics*. American Mathematical Soc., 2009.

K. Dykema and U. Haagerup (2004). DT-operators and decomposability of Voiculescu's circular operator. *American journal of mathematics*, 126(1):121–189.

P. J. Forrester. *Log-gases and random matrices*, volume 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2010.

P. J. Forrester, D. Wang, et al. (2017). Muttalib–Borodin ensembles in random matrix theory realisations and correlation functions. *Electronic Journal of Probability*, 22.

V. A. Marchenko and L. A. Pastur (1967). Distribution of eigenvalues for some sets of random matrices. *Matematicheskii Sbornik*, 114(4):507–536.

K. A. Muttalib (1995). Random matrix models with additional interactions. *Journal of Physics A: Mathematical and General*, 28(5):L159.

J. H. van Lint and R. M. Wilson. *A course in combinatorics*. Cambridge University Press, 2001.