

# CIRCULAR JACOBI ENSEMBLES AND DEFORMED VERBLUNSKY COEFFICIENTS

P. BOURGADE, A. NIKEGHBALI, AND A. ROUAULT

ABSTRACT. Using the spectral theory of unitary operators and the theory of orthogonal polynomials on the unit circle, we propose a simple matrix model for the following circular analogue of the Jacobi ensemble:

$$c_{\delta,\beta}^{(n)} \prod_{1 \leq k < l \leq n} |e^{i\theta_k} - e^{i\theta_l}|^\beta \prod_{j=1}^n (1 - e^{-i\theta_j})^\delta (1 - e^{i\theta_j})^{\bar{\delta}}$$

with  $\Re \delta > -1/2$ . If  $e$  is a cyclic vector for a unitary  $n \times n$  matrix  $U$ , the spectral measure of the pair  $(U, e)$  is well parameterized by its Verblunsky coefficients  $(\alpha_0, \dots, \alpha_{n-1})$ . We introduce here a deformation  $(\gamma_0, \dots, \gamma_{n-1})$  of these coefficients so that the associated Hessenberg matrix (called GGT) can be decomposed into a product  $r(\gamma_0) \cdots r(\gamma_{n-1})$  of elementary reflections parameterized by these coefficients. If  $\gamma_0, \dots, \gamma_{n-1}$  are independent random variables with some remarkable distributions, then the eigenvalues of the GGT matrix follow the circular Jacobi distribution above.

These deformed Verblunsky coefficients also allow to prove that, in the regime  $\delta = \delta(n)$  with  $\delta(n)/n \rightarrow \beta d/2$ , the spectral measure and the empirical spectral distribution weakly converge to an explicit nontrivial probability measure supported by an arc of the unit circle. We also prove the large deviations for the empirical spectral distribution.

## CONTENTS

1. Introduction	2
1.1. The circular Jacobi ensemble.	2
1.2. Orthogonal polynomials on the unit circle.	4
1.3. Organization of the paper.	8
2. Deformed Verblunsky coefficients and reflections	9
2.1. Analytical properties	9
2.2. Geometric interpretation	12
3. Deformed Verblunsky coefficients and independence	15
4. Matrix models for the circular Jacobi ensemble	19
5. Limiting spectral measure and large deviations	22
5.1. Weak convergence of measures	23

---

2000 *Mathematics Subject Classification.* 15A52, 60F05, 60F15, 60F10.

*Key words and phrases.* Random Matrices, circular Jacobi ensembles, Characteristic polynomial, Orthogonal polynomials on the unit circle, Verblunsky coefficients, spectral measure, Limit theorems.

5.2. Large deviations for the ESD	27
References	32

## 1. INTRODUCTION

**1.1. The circular Jacobi ensemble.** The theory of random unitary matrices was developed using the existence of a natural probability uniform measure on compact Lie groups, namely the Haar measure. The statistical properties of the eigenvalues as well as the characteristic polynomial of these random matrices have played a crucial role both in physics (see [27] for an historical account) and in analytic number theory to model  $L$ -functions (see [20] and [21] where Keating and Snaith predict moments of  $L$ -functions on the critical line using knowledge on the moments of the characteristic polynomial of random unitary matrices).

The circular unitary ensemble (CUE) is  $U(n)$ , the unitary group over  $\mathbb{C}^n$ , equipped with its Haar probability measure  $\mu_{U(n)}$ . Weyl's integration formula allows one to average any (bounded measurable) function on  $U(n)$  which is conjugation-invariant

$$\int f d\mu_{U(n)} = \frac{1}{n!} \int \cdots \int |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^2 f(\text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})) \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi}, \quad (1.1)$$

where  $\Delta(e^{i\theta_1}, \dots, e^{i\theta_n}) = \prod_{1 \leq j < k \leq n} (e^{i\theta_k} - e^{i\theta_j})$  denotes the Vandermonde determinant.

The circular orthogonal ensemble (COE) is the subset of  $U(n)$  consisting of symmetric matrices, i.e.  $U(n)/O(n) = \{VV^T; V \in U(n)\}$  equipped with the measure obtained by pushing forward  $\mu_{U(n)}$  by the mapping  $V \mapsto VV^T$ . The integration formula is similar to (1.1) but with  $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^2$  replaced by  $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|$  and with the normalizing constant changed accordingly.

For the circular symplectic ensemble (CSE), which will not be recalled here, the integration formula uses  $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^4$ .

Dyson observed that the induced eigenvalue distributions correspond to the Gibbs distribution for the classical Coulomb gas on the circle at three different temperatures. More generally,  $n$  identically charged particles confined to move on the unit circle, each interacting with the others through the usual Coulomb potential  $-\log|z_i - z_j|$ , give rise to the Gibbs measure at temperature  $1/\beta$  (see the discussion and references in [24] and in [13] chap. 2):

$$\mathbb{E}_n^\beta(f) = c_{0,\beta}^{(n)} \int f(e^{i\theta_1}, \dots, e^{i\theta_n}) |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta d\theta_1 \dots d\theta_n, \quad (1.2)$$

where  $c_{0,\beta}^{(n)}$  is a normalizing constant chosen so that

$$h_{0,\beta}^{(n)}(\theta_1, \dots, \theta_n) = c_{0,\beta}^{(n)} |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta \quad (1.3)$$

is a probability density on  $(0, 2\pi)^n$  and where  $f$  is any symmetric function. The unitary, orthogonal and symplectic circular ensembles correspond to matrix models for the Coulomb gas at  $\beta = 1, 2, 4$  respectively, but are there matrix models for general  $\beta > 0$  for Dyson's circular eigenvalue statistics?

Killip and Nenciu [24] provided matrix models for Dyson's circular ensemble, using the theory of orthogonal polynomials on the unit circle. In particular, they obtained a sparse matrix model which is five-diagonal, called CMV (after the names of the authors Cantero, Moral, Velásquez [9]). In this framework, there is not a natural underlying measure such as the Haar measure; the matrix ensemble is characterized by the laws of its elements.

There is an analogue of Dyson's circular ensembles on the real line: the probability density function of the eigenvalues  $(x_1, \dots, x_n)$  for such ensembles with temperature  $1/\beta$  is proportional to

$$|\Delta(x_1, \dots, x_n)|^\beta \prod_{j=1}^n e^{-x_j^2/2} \quad (1.4)$$

For  $\beta = 1, 2$  or  $4$ , this corresponds to the classical Gaussian ensembles. Dimitriu and Edelman [11] gave a simple tridiagonal matrix model for (1.4). Killip and Nenciu [24], gave an analogue matrix model for the Jacobi measure on  $[-2, 2]^n$ , which is up to a normalizing constant,

$$|\Delta(x_1, \dots, x_n)|^\beta \prod_{j=1}^n (2 - x_j)^a (2 + x_j)^b dx_1 \dots dx_n, \quad (1.5)$$

where  $a, b > 0$ , relying on the theory of orthogonal polynomials on the unit circle and its links with orthogonal polynomials on the segment. When  $a$  and  $b$  are strictly positive integers, the Jacobi measure (1.5) can be interpreted as the potential  $|\Delta(x_1, \dots, x_{n+a+b})|^\beta$  on  $[-2, 2]^{n+a+b}$  conditioned to have  $a$  elements located at  $2$  and  $b$  elements located at  $-2$ . Consequently, the Jacobi measure on the unit circle should be a two parameters extension of (1.3), corresponding to conditioning to have specific given eigenvalues. Such an analogue was defined as the *circular Jacobi ensemble* in [13] and [15]. If  $\delta \in \mathbb{R}$ , we recover the cJUE as in Witte and Forrester [37].

**Definition 1.1.** Throughout this paper, we note  $h_{\delta, \beta}^{(n)}$  the probability density function on  $(0, 2\pi)$  given by:

$$h_{\delta, \beta}^{(n)}(\theta_1, \dots, \theta_n) = c_{\delta, \beta}^{(n)} |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta \prod_{j=1}^n (1 - e^{-i\theta_j})^\delta (1 - e^{i\theta_j})^{\bar{\delta}} \quad (1.6)$$

with  $\delta \in \mathbb{C}$ ,  $\Re(\delta) > -\frac{1}{2}$ .

If  $\delta \in \frac{\beta}{2}\mathbb{N}$ , this measure coincides with (1.3) conditioned to have eigenvalues at  $1$ . For  $\beta = 2$ , such measures were first considered by Hua [19] and Pickrell [30], [31]. This case was also widely studied in [29] and [6] for its

connections with the theory of representations and in [8] for its analogies with the Ewens measures on permutation groups.

One of our goals in this paper is to provide matrix models for the circular Jacobi ensemble, i.e. a distribution on  $U(n)$  such that the arguments of the eigenvalues  $(e^{i\theta_1}, \dots, e^{i\theta_n})$  are distributed as in (1.6). One can guess that additional problems may appear because the distribution of the eigenvalues is not rotation invariant anymore. Nevertheless, some statistical information for the circular Jacobi ensemble can be obtained from Dyson's circular ensemble by a sampling (or a change of probability measure) with the help of the determinant. More precisely, let us first define the notion of sampling.

**Definition 1.2.** Let  $(X, \mathcal{F}, \mu)$  be a probability space, and  $h : X \mapsto \mathbb{R}^+$  a measurable and integrable function with  $\mathbb{E}_\mu(h) > 0$ . Then a measure  $\mu'$  is said to be the  $h$ -sampling of  $\mu$  if for all bounded measurable functions  $f$

$$\mathbb{E}_{\mu'}(f) = \frac{\mathbb{E}_\mu(fh)}{\mathbb{E}_\mu(h)}.$$

If we consider a matrix model for  $h_{0,\beta}^{(n)}$ , we can define<sup>1</sup> a matrix model for  $h_{\delta,\beta}^{(n)}$  by the means of a sampling, noticing that when the charges are actually the eigenvalues of a matrix  $U$ , then (1.3) differs from (1.6) by a factor which is a function of  $\det(\text{Id} - U)$ . We define  $\det_\delta$  for a unitary matrix  $U$  as

$$\det_\delta(U) = \det(\text{Id} - U)^\delta \det(\text{Id} - \bar{U})^\delta,$$

and we will use this  $\det_\delta$ -sampling.

Actually we look for an effective construction of a random matrix, for instance starting from a reduced number of independent random variables with known distributions. Notice that in the particular case  $\beta = 2$ , the density  $h_{0,2}$  corresponds to eigenvalues of a matrix under the Haar measure on  $U(n)$  and the  $\det_\delta$ -sampling of this measure is the Hua-Pickrell measure studied in our previous paper [8].

**1.2. Orthogonal polynomials on the unit circle.** We now wish to outline the main ingredients which are needed from the theory of orthogonal polynomials on the unit circle to construct matrix models for the general Dyson's circular ensemble. The reader can refer to [35] and [36] for more results and references; in particular, all the results about orthogonal polynomials on the unit circle (named hereafter OPUC) can be found in these volumes.

Let us explain why OPUC play a prominent role in these constructions. In all this paper,  $\mathbb{D}$  denotes the open unit disk  $\{z \in \mathbb{C} : |z| < 1\}$  and  $\mathbb{T}$  the unit circle  $\{z \in \mathbb{C} : |z| = 1\}$ . Let  $(\mathcal{H}, u, e)$  be a triple where  $\mathcal{H}$  is a Hilbert space,  $u$  a unitary operator and  $e$  a cyclic unit vector, i.e.  $\{u^j e\}_{j=-\infty}^\infty$  is total in  $\mathcal{H}$ . We say that two triples  $(\mathcal{H}, u, e)$  and  $(\mathcal{K}, v, e')$  are equivalent if and only if there exists an isometry  $k : \mathcal{H} \rightarrow \mathcal{K}$  such that  $v = k u k^{-1}$  and

<sup>1</sup>for  $\Re(\delta) > -1/2$ , due to an integrability constraint

$e' = ke$ . The spectral theorem says that for each equivalence class, there exists a unique probability measure  $\mu$  on  $\mathbb{T}$  such that

$$\langle e, u^k e \rangle_{\mathcal{H}} = \int_{\mathbb{T}} z^k d\mu(z) \quad , \quad k = 0, \pm 1, \dots$$

Conversely, such a probability measure  $\mu$  gives rise to a triple consisting of the Hilbert space  $L^2(\mathbb{T}, \mu)$ , the operator of multiplication by  $z$ , i.e.  $h \mapsto (z \mapsto zh(z))$  and the vector  $\mathbf{1}$ , i.e. the constant function 1. When the space  $\mathcal{H}$  is fixed, the probability measure  $\mu$  associated with the triple  $(\mathcal{H}, u, e)$  is called the *spectral measure of the pair*  $(u, e)$ .

Let us consider the finite  $n$ -dimensional case. Assume that  $u$  is unitary and  $e$  is cyclic. It is classical that  $u$  has  $n$  different eigenvalues  $(e^{i\theta_j}, j = 1, \dots, n)$  with  $\theta_j \in [0, 2\pi)$ . In any orthonormal basis whose first vector is  $e$ , say  $(e_1 = e, e_2, \dots, e_n)$ ,  $u$  is represented by a matrix  $U$  and there is a unitary matrix  $\Pi$  diagonalizing  $U$ . It is then straightforward that the spectral measure is

$$\mu = \sum_{j=1}^n \pi_j \delta_{e^{i\theta_j}} \quad (1.7)$$

where the weights are defined as  $\pi_j = |\langle e_1, \Pi e_j \rangle|^2$ . Note that  $\pi_j > 0$  because a cyclic vector cannot be orthogonal to any eigenvector (and we also have  $\sum_{j=1}^n \pi_j = 1$  because  $\Pi$  is unitary). The eigenvalues  $(e^{i\theta_j}, j = 1, \dots, n)$  and the vector  $(\pi_1, \dots, \pi_n)$  can then be used as coordinates for the probability measure  $\mu$ .

Keeping in mind our purpose, we see that the construction of a matrix model from a vector  $(e^{i\theta_j}, j = 1, \dots, n)$  may be achieved in two steps: first give a vector of weights  $(\pi_1, \dots, \pi_n)$ , then find a matricial representative of the equivalence class with a rather simple form. The key tool for the second task is the sequence of orthogonal polynomials associated with the measure  $\mu$ . In  $L^2(\mathbb{T}, \mu)$  equipped with the natural basis  $\{1, z, z^2, \dots, z^{n-1}\}$ , the Gram-Schmidt procedure provides the family of monic orthogonal polynomials  $\Phi_0, \dots, \Phi_{n-1}$ . We can still define  $\Phi_n$  as the unique monic polynomial of degree  $n$  with  $\|\Phi_n\|_{L^2(\mathbb{T}, \mu)} = 0$ , namely

$$\Phi_n(z) = \prod_{j=1}^n (z - e^{i\theta_j}). \quad (1.8)$$

The  $\Phi_k$ 's ( $k = 0, \dots, n$ ) obey the Szegő recursion relation:

$$\Phi_{j+1}(z) = z\Phi_j(z) - \bar{\alpha}_j \Phi_j^*(z) \quad (1.9)$$

where

$$\Phi_j^*(z) = z^j \overline{\Phi_j(\bar{z}^{-1})}. \quad (1.10)$$

The coefficients  $\alpha_j$ 's ( $0 \leq j \leq n-1$ ) are called Verblunsky coefficients and satisfy the condition  $\alpha_0, \dots, \alpha_{n-2} \in \mathbb{D}$  and  $\alpha_{n-1} \in \mathbb{T}$ .

When the measure  $\mu$  has infinite support, one can define the family of orthogonal polynomials  $(\Phi_n)_{n \geq 0}$  associated with  $\mu$  for all  $n$ . Then there are infinitely many Verblunsky coefficients  $(\alpha_n)$  which all lie in  $\mathbb{D}$ .

Verblunsky's Theorem (see for example [35], [36]) states that there is a bijection between probability measures on the unit circle and sequences of Verblunsky coefficients. The matrix of the multiplication by  $z$  in  $L^2(\mathbb{T}, \mu)$ , in the basis of orthonormal polynomials, has received much attention. This unitary matrix, noted  $\mathcal{G}(\alpha_0, \dots, \alpha_{n-1})$ , called GGT by B. Simon (see Chapter 4.1 in [35] for more details and for an historical account), is in the Hessenberg form: all entries below the subdiagonal are zero, whereas the entries above the subdiagonal are nonzero and the subdiagonal is nonnegative (see formulae (4.1.5) and (4.1.6) in [35] for an explicit expression for the entries in terms of the Verblunsky coefficients, or formula (1.11) in Lemma 1.3 below).

For  $H$  a  $n \times n$  complex matrix, the subscript  $H_{ij}$  stands for  $\langle e_i, H(e_j) \rangle$ , where  $\langle x, y \rangle = \sum_{k=1}^n \bar{x}_k y_k$ . Killip and Nenciu state that any unitary matrix in the Hessenberg form with nonnegative subdiagonal is the matrix of multiplication by  $z$  in  $L^2(\mathbb{T}, \mu)$  for some measure  $\mu$ . More precisely, from Killip-Nenciu [24], Lemma 3.2, we have

**Lemma 1.3.** *Let  $\mu$  be a probability measure on  $\mathbb{T}$  supported at  $n$  points. Then the matrix  $H$  of  $f(z) \mapsto zf(z)$  in the basis of orthonormal polynomials of  $L^2(\mathbb{T}, \mu)$ , is in the Hessenberg form. More precisely*

$$H_{i+1, j+1} = \begin{cases} -\alpha_{i-1} \bar{\alpha}_j \prod_{p=i}^{j-1} \rho_p & \text{if } i < j + 1 \\ \rho_{j-1} & \text{if } i = j + 1, \\ 0 & \text{if } i > j + 1 \end{cases}, \quad (1.11)$$

with  $\rho_j = \sqrt{1 - |\alpha_j|^2}$  and  $\alpha_{-1} = -1$ , the  $\alpha_k$ 's being the Verblunsky coefficients associated to  $\mu$ . Conversely, if  $\alpha_0, \dots, \alpha_{n-1}$  are given in  $\mathbb{D}^{n-1} \times \mathbb{T}$ , and if we define the matrix  $H$  by (1.11), then the spectral measure of the pair  $(H, e_1)$  is the measure  $\mu$  whose Verblunsky coefficients are precisely  $\alpha_0, \dots, \alpha_{n-1}$ .

Besides, there is a very useful decomposition of these matrices into product of block matrices, called the AGR decomposition by Simon ([33]), after the paper [3]. For  $0 \leq k \leq n - 2$ , let

$$\Theta^{(k)}(\alpha) = \text{Id}_k \oplus \begin{pmatrix} \bar{\alpha}_k & \rho_k \\ \rho_k & -\alpha_k \end{pmatrix} \oplus \text{Id}_{n-k-2}.$$

and set  $\Theta^{(n-1)}(\alpha_{n-1}) = \text{Id}_{n-1} \oplus (\bar{\alpha}_{n-1})$ , with  $|\alpha_{n-1}| = 1$ . Then the AGR decomposition states that ([33] Theorem 10.1)

$$\mathcal{G}(\alpha_0, \dots, \alpha_{n-1}) = \Theta^{(0)}(\alpha_0) \Theta^{(1)}(\alpha_1) \dots \Theta^{(n-1)}(\alpha_{n-1}).$$

Now we state a crucial result of Killip and Nenciu which enabled them to obtain a matrix model in the Hessenberg form for Dyson's circular ensemble.

The challenge consists in randomizing  $(\alpha_0, \dots, \alpha_{n-1})$  in  $\mathbb{D}^{n-1} \times \mathbb{T}$  in such a way that the angles  $(\theta_1, \dots, \theta_n)$  of the spectral measure

$$\mu = \sum_{j=1}^n \pi_j \delta_{e^{i\theta_j}}$$

have the density  $h_{0,\beta}^{(\delta)}$  (see (1.6)). To make the statement precise, let us introduce three definitions (for the properties of the Beta and Dirichlet distributions, see [32]).

**Definition 1.4.** For  $a_1, a_2 > 0$ , let  $\text{Beta}(a_1, a_2)$  be the distribution on  $[0, 1]$  with density

$$\frac{\Gamma(a_1 + a_2)}{\Gamma(a_1)\Gamma(a_2)} x^{a_1-1} (1-x)^{a_2-1}$$

Its generalization is the following.

**Definition 1.5.** For  $n \geq 2$  and  $a_1, \dots, a_n > 0$ , let  $\text{Dir}(a_1, \dots, a_n)$  be the distribution on the simplex  $\{(x_1, \dots, x_n) \in [0, 1]^n : \sum_{i=1}^n x_i = 1\}$  with density

$$\frac{\Gamma(a_1 + \dots + a_n)}{\Gamma(a_1) \cdots \Gamma(a_n)} \prod_{k=1}^n x_k^{a_k-1}.$$

If  $a_1 = \dots = a_n = a$ , it is called the Dirichlet distribution of order  $n \geq 2$  with parameter  $a > 0$  and denoted by  $\text{Dir}_n(a)$ . (For  $a = 1$ , this is the uniform distribution).

**Definition 1.6.** For  $s > 1$  let  $\nu_s$  be the probability measure on  $\mathbb{D}$  with density

$$\frac{s-1}{2\pi} (1-|z|^2)^{(s-3)/2}.$$

It is the law of  $\mathbb{R}e^{i\psi}$  where  $\mathbb{R}$  and  $\psi$  are independent,  $\psi$  is uniformly distributed on  $(0, 2\pi)$  and  $\mathbb{R}^2$  has the  $\text{Beta}(1, (s-1)/2)$  distribution. We adopt the convention that  $\nu_1$  is the uniform distribution on the unit circle. We denote by  $\eta_{0,\beta}^{(n)}$  the distribution on  $\mathbb{D}^{n-1} \times \mathbb{T}$  given by

$$\eta_{0,\beta}^{(n)} = \otimes_{k=0}^{n-1} \nu_{\beta(n-k-1)+1}. \quad (1.12)$$

**Proposition 1.7** (Killip-Nenciu [24], Proposition 4.2). *The following formulae express the same measure on the manifold of probability distributions on  $\mathbb{T}$  supported at  $n$  points:*

$$\frac{2^{1-n}}{n!} |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta \prod_{j=1}^n \pi_j^{\beta/2-1} d\theta_1 \dots d\theta_n d\pi_1 \dots d\pi_{n-1}$$

in the  $(\theta, \pi)$  coordinates and

$$\prod_{k=0}^{n-2} (1-|\alpha_k|^2)^{(\beta/2)(n-k-1)-1} d^2\alpha_0 \dots d^2\alpha_{n-2} \frac{d\phi}{2\pi} \quad (1.13)$$

in terms of the Verblunsky coefficients.

Proposition 1.7 may be restated as follows: to pick at random a measure  $\mu$  such that  $(\alpha_0, \dots, \alpha_{n-1})$  is  $\eta_{0,\beta}^{(n)}$  distributed is equivalent to pick the support  $(\theta_1, \dots, \theta_n)$  according to  $h_{0,\beta}^{(n)}$  (see (1.3)) and to pick the weights  $(\pi_1, \dots, \pi_n)$  independently according to  $\text{Dir}_n(\beta/2)$ .

As a consequence, if one takes independent coefficients  $(\alpha_0, \dots, \alpha_{n-1})$  such that  $\alpha_k$  is  $\nu_{\beta(n-k-1)+1}$  distributed for  $0 \leq k \leq n-1$ , then the GGT matrix  $\mathcal{G}(\alpha_0, \dots, \alpha_{n-1})$  will be a matrix model for Dyson's circular ensemble with temperature  $1/\beta$  (see also Proposition 2.11 in [13]). Actually in [24], Killip and Nenciu provide a matrix model which is much sparser (five-diagonal) as shall be explained in Section 4.

Let us now define the laws on  $U(n)$  which we will consider in the sequel.

**Definition 1.8.** We denote by  $\text{CJ}_{0,\beta}^{(n)}$  the probability distribution supported by the set of  $n \times n$  GGT matrices of the form (1.11), corresponding to the law of  $\mathcal{G}(\alpha_0, \dots, \alpha_{n-1})$  defined above. We denote by  $\text{CJ}_{\delta,\beta}^{(n)}$  the probability distribution on  $U(n)$  which is the  $\det_\delta$ -sampling of  $\text{CJ}_{0,\beta}^{(n)}$ .

The standard GGT approach is not sufficient to produce matrix models for the circular Jacobi ensemble because, as we shall see in Section 3, under the measure  $\text{CJ}_{\delta,\beta}^{(n)}$ , the Verblunsky coefficients are not independent anymore. To overcome this difficulty, we associate to a measure on the unit circle, or equivalently to its Verblunsky coefficients, a new sequence of coefficients  $(\gamma_k)_{0 \leq k \leq n-1}$ , which we shall call deformed Verblunsky coefficients. There is a simple bijection between the original sequence  $(\alpha_k)_{0 \leq k \leq n-1}$  and the new one  $(\gamma_k)_{0 \leq k \leq n-1}$ . These coefficients satisfy among other nice properties that  $|\alpha_k| = |\gamma_k|$ , and that they are independent under  $\text{CJ}_{\delta,\beta}^{(n)}$  (and for  $\delta = 0$  the  $\alpha_k$ 's and the  $\gamma_k$ 's have the same distribution). They have a geometric interpretation in terms of reflections: this leads to a decomposition of the GGT matrix  $\mathcal{G}(\alpha_0, \dots, \alpha_{n-1})$  as a product of independent elementary reflections (constructed from the  $\gamma_k$ 's). The explicit expression of the densities allows an asymptotic study (as  $n \rightarrow \infty$ ) of the  $\gamma_k$ 's, and consequently of the spectral measure, and finally of the empirical spectral distribution.

**1.3. Organization of the paper.** In Section 2, after recalling basic facts about the reflections introduced in [8], we define the deformed Verblunsky coefficients  $(\gamma_k)_{0 \leq k \leq n-1}$  and give some of its basic properties. In particular we prove that the GGT matrix  $\mathcal{G}(\alpha_0, \dots, \alpha_{n-1})$  can be decomposed into a product of elementary complex reflections (Theorem 2.8).

In Section 3, we derive the law of the  $\gamma_k$ 's under  $\text{CJ}_{\delta,\beta}^{(n)}$  (Theorem 3.2); in particular we show that they are independent and that the actual Verblunsky coefficients are dependent if  $\delta \neq 0$ . We then prove an analogue of the above Proposition 1.7 on the  $(\theta, \pi)$  coordinates of  $\mu$  (Theorem 3.3).



In Section 4, we propose our matrix model (Theorem 4.1). It is a modification of the AGR factorization, where we transform the  $\Theta_k$ 's so that they become reflections :

$$\Xi^{(k)}(\alpha) = \text{Id}_k \oplus \begin{pmatrix} \bar{\alpha} & e^{i\phi}\rho \\ \rho & -e^{i\phi}\alpha \end{pmatrix} \oplus \text{Id}_{n-k-2},$$

with  $e^{i\phi} = \frac{1-\bar{\alpha}}{1-\alpha}$ . Of course the CMV representation [9], which is five-diagonal, is also available, but this time the  $\alpha_k$ 's are not independent. Using the following elementary fact proven in Section 2,

$$\Phi_n(1) = \det(\text{Id} - U) = \prod_{k=0}^{n-1} (1 - \gamma_k),$$

we are able to generalize our previous results in [7] and [8] about the decomposition of the characteristic polynomial evaluated at 1 as a product of independent complex variables (Proposition 4.3).

In Section 5, we study asymptotic properties of our model as  $n \rightarrow \infty$ , when  $\delta = \beta n d / 2$ , with  $\Re d \geq 0$ . We first prove that the Verblunsky coefficients have deterministic limits in probability. This entails that the spectral measure converges weakly in probability to the same deterministic measure (denoted by  $\mu_d$ ) which is supported by an arc of the unit circle (Theorem 5.1). Besides, we consider the empirical spectral distribution (ESD), where the Dirac masses have the same weight  $1/n$ . Bounding the distances between both random measures, we prove that the ESD has the same limit (Theorem 5.4). Moreover, starting from the explicit joint distribution (1.6), we prove also that the ESD satisfies the large deviation principle at scale  $(\beta/2)n^2$  whose rate function reaches its minimum at  $\mu_d$  (Theorem 5.5).

## 2. DEFORMED VERBLUNSKY COEFFICIENTS AND REFLECTIONS

In this section, we introduce the deformed Verblunsky coefficients and we establish some of their relevant properties, in particular a geometric interpretation in terms of reflections. One remarkable property of the Verblunsky coefficients, as it appears in Proposition 1.7, is that they are independent under  $\text{CJ}_{0,\beta}^{(n)}$ . As we shall see in Section 3, this does not hold anymore under  $\text{CJ}_{\delta,\beta}^{(n)}$ . This motivated us to introduce a new set of coefficients,  $(\gamma_0, \dots, \gamma_{n-2}, \gamma_{n-1})$ , called deformed Verblunsky coefficients, which are uniquely associated with a set of Verblunsky coefficients. In particular,  $\gamma_k \in \mathbb{D}$  for  $0 \leq k \leq n-2$ ,  $\gamma_{n-1} \in \mathbb{T}$  and the map  $(\gamma_0, \dots, \gamma_{n-1}) \mapsto (\alpha_0, \dots, \alpha_{n-1})$  is a bijection. Moreover, the characteristic polynomial at 1 can be expressed simply in terms of  $(\gamma_0, \dots, \gamma_{n-1})$ .

**2.1. Analytical properties.** Let  $\mu$  be a probability measure on the unit circle supported at  $n$  points. Keeping the notations of the introduction, we let  $(\Phi_k(z))_{0 \leq k \leq n}$  denote the monic orthogonal polynomials associated with

$\mu$  and  $(\alpha_k)_{0 \leq k \leq n-1}$  its corresponding set of Verblunsky coefficients through Szegő's recursion formula (1.9). The functions

$$b_k(z) = \frac{\Phi_k(z)}{\Phi_k^*(z)}, \quad k \leq n-1 \quad (2.1)$$

are known as the inverse Schur iterates ([36] p.476, after Khrushchev [23] p.273). They are analytic in a neighborhood of  $\mathbb{D}$  and meromorphic in  $\mathbb{C}$ . Each  $b_k$  is a finite Blaschke product

$$b_k(z) = \prod_{j=1}^n \left( \frac{z - z_j}{1 - \bar{z}_j z} \right)$$

where  $z_1, \dots, z_k$  are the zeros of  $\Phi_k$ . Let us now explain the term "inverse Schur iterate".

The Schur function is a fundamental object in the study of the orthogonal polynomials on the unit circle. Let us briefly recall its definition (see [35] or [34] for more details and proofs): if  $\mu$  is a probability measure on the unit circle (supported at finitely many points or not), its Schur function  $f : \mathbb{D} \rightarrow \mathbb{D}$  is defined as:

$$f(z) = \frac{1}{z} \frac{F(z) - 1}{F(z) + 1} \quad \text{where} \quad F(z) = \int \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu(e^{i\theta}). \quad (2.2)$$

It is a bijection between the set of probability measures on the unit circle and analytic functions mapping  $\mathbb{D}$  to  $\mathbb{D}$ . The Schur algorithm (which is described in [35] or [34] p.438) allows to parametrize the Schur function  $f$  by a sequence of so-called Schur parameters, which are actually the Verblunsky coefficients associated to  $\mu$  (Geronimus theorem). In particular, there are finitely many Verblunsky coefficients (or equivalently the measure  $\mu$  is supported at  $n$  points) if and only if  $f$  is a finite Blaschke product. The name "inverse Schur iterate" ([22]) for  $b_k$  comes from the result (1.12) of the latter paper where  $b_k$  is identified as the Schur function corresponding to the "reversed sequence"  $(-\bar{\alpha}_{k-1}, \dots, -\bar{\alpha}_0, 1)$  (see also [36] Prop. 9.2.3).

Let us define our sequence of functions, which shall lead us the deformed coefficients.

**Definition 2.1.** If  $\mu$  is supported at  $n$  points and with the notation above, define  $\gamma_k(z)$  for  $0 \leq k \leq n-1$ , as:

$$\gamma_k(z) = z - \frac{\Phi_{k+1}(z)}{\Phi_k(z)}. \quad (2.3)$$

From the Szegő's recursion formula (1.9) and notation (2.1), this is equivalent to

$$\gamma_k(z) = \frac{\bar{\alpha}_k}{b_k(z)}, \quad (2.4)$$

so that  $\gamma_k$  is meromorphic, with poles in  $\mathbb{D}$  and zeros lying outside  $\bar{\mathbb{D}}$ .

The next proposition shows how the functions  $\gamma_k(z)$  can be defined recursively with the help of the coefficients  $\alpha_k$ . As a consequence, we shall see that the  $\gamma_k(z)$  are very closely related to a fundamental object in the theory of random matrices: the characteristic polynomial.

**Proposition 2.2.** *For any  $z \in \mathbb{C}$ ,  $\gamma_0(z) = \bar{\alpha}_0$  and the following decomposition for  $\Phi_k(z)$  holds:*

$$\Phi_k(z) = \prod_{j=0}^{k-1} (z - \gamma_j(z)), \quad k = 1, \dots, n. \quad (2.5)$$

The  $\gamma_k(z)$ 's may be also defined by means of the  $\alpha$ 's through the recursion :

$$\gamma_k(z) = \bar{\alpha}_k \prod_{j=0}^{k-1} \frac{1 - z\tilde{\gamma}_j(z)}{z - \gamma_j(z)}, \quad (2.6)$$

$$\tilde{\gamma}_k(z) = \overline{\gamma_k(\bar{z}^{-1})}. \quad (2.7)$$

*Proof.* The first claim is an immediate consequence of (2.3). Now, using  $\Phi_k(z) = \prod_{j=0}^{k-1} (z - \gamma_j(z))$ , we obtain

$$\Phi_k^*(z) = \prod_{j=0}^{k-1} (1 - z\tilde{\gamma}_j(z)),$$

and hence (we use (2.4))

$$\gamma_k(z) = \bar{\alpha}_k \prod_{j=0}^{k-1} \frac{1 - z\tilde{\gamma}_j(z)}{z - \gamma_j(z)}.$$

□

Note that when  $|z| = 1$ ,  $|\gamma_k(z)| = |\alpha_k|$ . Combined with the above proposition, this leads us to introduce the following set of coefficients.

**Definition 2.3.** Define the coefficients  $(\gamma_k)_{0 \leq k \leq n-1}$  by

$$\gamma_k := \gamma_k(1), \quad k = 0, \dots, n-1. \quad (2.8)$$

We shall refer to the  $\gamma_k$ 's as the *deformed Verblunsky coefficients*.

**Proposition 2.4.** *The following properties hold for the deformed Verblunsky coefficients:*

- a) For all  $0 \leq k \leq n-1$ ,  $|\gamma_k| = |\alpha_k|$ , and in particular  $\gamma_{n-1} \in \mathbb{T}$ ;
- b)  $\gamma_0 = \bar{\alpha}_0$  and

$$\gamma_k = \bar{\alpha}_k e^{i\varphi_{k-1}}, \quad e^{i\varphi_{k-1}} = \prod_{j=0}^{k-1} \frac{1 - \tilde{\gamma}_j}{1 - \gamma_j}, \quad (k = 1, \dots, n-1). \quad (2.9)$$

The last term is special. Since  $|\alpha_{n-1}| = 1$ , we set  $\alpha_{n-1} = e^{i\psi_{n-1}}$ , so that

$$\gamma_{n-1} = e^{i(-\psi_{n-1} + \varphi_{n-2})} := e^{i\theta_{n-1}}. \quad (2.10)$$

c) Let  $\mu$  be the spectral measure associated to  $(U, e_1)$ ,  $U \in U(n)$ . Then  $\Phi_n(z)$  is the characteristic polynomial of  $U$ , in particular,

$$\Phi_n(1) = \det(\text{Id} - U) = \prod_{k=0}^{n-1} (1 - \gamma_k). \quad (2.11)$$

*Proof.* All the results are direct consequences of the definition 2.3 and the formulae in Proposition 2.2 evaluated at 1.  $\square$

*Remark.* In [26], Killip and Stoiciu have already considered variables which are the complex conjugate of our deformed Verblunsky coefficients as auxiliary variables in the study of the Prüfer phase (Lemma 2.1 in [26]). Nevertheless, the way we define them as well as the use we make of them are different.

*Remark.* The formula (2.9) shows that the  $\gamma_k$ 's can be obtained from the  $\alpha_k$ 's recursively. Hence starting from a spectral measure associated to a unitary matrix, one can associate with it the Verblunsky coefficients and then the deformed Verblunsky coefficients. Conversely, one can translate any property of the deformed ones into properties for the spectral measure associated with it by inverting the transformations (2.9).

*Remark.* The distribution of the characteristic polynomial of random unitary matrices evaluated at 1, through its Mellin-Fourier transform, plays a key role in the theory of random matrices, especially through its links with analytic number theory (see [28] for an account). In [7] it is proven that it can be decomposed in law into a product of independent random variables when working on the unitary and orthogonal groups endowed with the Haar measure; since we will prove in Section 3 that the  $\gamma_k$ 's are independent under  $\text{CJ}_{\delta, \beta}^{(n)}$ , then we can conclude that this latter result holds for any circular Jacobi ensemble.

**2.2. Geometric interpretation.** We give a connection between the coefficients  $(\gamma_k)_{0 \leq k \leq n-1}$  and reflections defined just below. This allows us to obtain a new decomposition of the GGT matrix associated with a measure  $\mu$  supported at  $n$  points on the unit circle as a product of  $n$  elementary reflections.

Many distinct definitions of reflections on the unitary group exist, the most well-known may be the Householder reflections. The transformations which will be relevant to us are the following ones.

**Definition 2.5.** An element  $r$  in  $U(n)$  will be referred to as a reflection if  $r - \text{Id}$  has rank 0 or 1.

If  $v \in \mathbb{C}^n$ , we denote by  $\langle v |$  the linear form  $w \mapsto \langle v, w \rangle$ . The reflections can also be described in the following way. If  $e$  and  $m \neq e$  are unit vectors of  $\mathbb{C}^n$ , there is a unique reflection  $r$  such that  $r(e) = m$ , and

$$r = \text{Id} - \frac{1}{1 - \langle m, e \rangle} (m - e) \langle (m - e) |. \quad (2.12)$$

Let  $F := \text{span}\{e, m\}$  be the 2-dimensional vector space which is spanned by the vectors  $e$  and  $m$ . It is clear that the reflection given by formula (2.12) leaves  $F^\perp$  invariant. Now set

$$\gamma = \langle e, m \rangle, \quad \rho = \sqrt{1 - |\gamma|^2}, \quad e^{i\varphi} = \frac{1 - \gamma}{1 - \bar{\gamma}}, \quad (2.13)$$

and let  $g \in F$  be the unit vector orthogonal to  $e$  obtained by the Gram-Schmidt procedure. Then in the basis  $(e, g)$  of  $F$ , the matrix of the restriction of  $r$  is

$$\Xi(\gamma) := \begin{pmatrix} \gamma & \rho e^{i\varphi} \\ \rho & -\bar{\gamma} e^{i\varphi} \end{pmatrix}. \quad (2.14)$$

Conversely, for  $\gamma \in \mathbb{D}$ , such a matrix represents the unique reflection in  $\mathbb{C}^2$  provided with its canonical basis, mapping  $e_1$  onto  $\gamma e_1 + \sqrt{1 - |\gamma|^2} e_2$ . The eigenvalues of  $r$  are 1 and  $-e^{i\varphi}$ .

Let  $u$  be a unitary operator in  $\mathbb{C}^n$  and  $e$  a unit cyclic vector for  $u$ . We define  $n$  reflections  $r_1, \dots, r_n$  recursively as follows. Let  $(\varepsilon_1, \dots, \varepsilon_n)$  be the orthonormal basis obtained from the Gram-Schmidt procedure applied to  $(e, ue, \dots, u^{n-1}e)$ .

Let  $r_1$  be the reflection, mapping  $e = \varepsilon_1$  onto  $ue = u\varepsilon_1$ . More generally, for  $k \geq 2$  let  $r_k$  be the reflection mapping  $\varepsilon_k$  onto  $r_{k-1}^{-1} r_{k-2}^{-1} \dots r_1^{-1} u\varepsilon_k$ . We will identify these reflections and establish the decomposition of  $u$ . Following the basics recalled about GGT matrices in the introduction, we note that the matrix of  $u$  in the basis  $(\varepsilon_1, \dots, \varepsilon_n)$  is the GGT matrix associated to the measure  $\mu$ , i.e. the matrix  $\mathcal{G}(\alpha_0, \dots, \alpha_{n-2}, \alpha_{n-1})$ , where  $(\alpha_0, \dots, \alpha_{n-2}, \alpha_{n-1})$  are the Verblunsky coefficients associated with the measure  $\mu$ . We will use formula (4.1.6) of [35] or formula (1.11) of (our) Lemma 1.3 for the identification of scalar products.

**Proposition 2.6.** *(1) For every  $1 \leq k \leq n - 1$ , the reflection  $r_k$  leaves invariant the  $n - 2$ -dimensional space  $\text{Span}\{\varepsilon_1, \dots, \varepsilon_{k-1}, \varepsilon_{k+2}, \dots, \varepsilon_n\}$ . The reflection  $r_n$  leaves invariant  $\text{Span}\{\varepsilon_1, \dots, \varepsilon_{n-1}\}$ .  
 (2) The following decomposition holds :*

$$u = r_1 \cdots r_n. \quad (2.15)$$

*Proof.* (1) In view of Section 2.1, it is enough to prove that for  $j \notin \{k, k+1\}$ , the vectors  $\varepsilon_j$  and  $r_k \varepsilon_k$  are orthogonal.

For  $k = 1$ ,  $\langle \varepsilon_j, r_1 \varepsilon_1 \rangle = \langle \varepsilon_j, u\varepsilon_1 \rangle = 0$  as soon as  $j \geq 3$  from (1.11).

Assume that for every  $\ell \leq k - 1$ , the reflection  $r_\ell$  leaves invariant  $\text{Span}\{\varepsilon_1, \dots, \varepsilon_{\ell-1}, \varepsilon_{\ell+2}, \dots, \varepsilon_n\}$ . For every  $j = 1, \dots, n$ , we have

$$\langle \varepsilon_j, r_k \varepsilon_k \rangle = \langle \varepsilon_j, r_{k-1}^{-1} r_{k-2}^{-1} \dots r_1^{-1} u\varepsilon_k \rangle = \langle r_1 \cdots r_{k-1} \varepsilon_j, u\varepsilon_k \rangle. \quad (2.16)$$

For  $j \geq k + 2$ , by assumption, the reflections  $r_1, \dots, r_{k-1}$  leave invariant  $\varepsilon_j$ , so that the above expression reduces to  $\langle \varepsilon_j, u\varepsilon_k \rangle$  which is 0 again by (1.11).

For  $j = k - 1$ , we have  $r_1 \cdots r_{k-1} \varepsilon_{k-1} = u\varepsilon_{k-1}$  by definition of  $r_{k-1}$ , so that (2.16) gives  $\langle \varepsilon_{k-1}, r_k \varepsilon_k \rangle = \langle u\varepsilon_{k-1}, u\varepsilon_k \rangle$ , which is 0 since  $u$  is unitary.

For  $j < k-1$ , by assumption, the reflections  $r_{j+1}, \dots, r_{k-1}$  leave invariant  $\varepsilon_j$ , so that the right hand side of (2.16) reduces to  $\langle r_1 \cdots r_j \varepsilon_j, u\varepsilon_k \rangle$ . By definition of  $r_j$ , it is  $\langle u\varepsilon_j, u\varepsilon_k \rangle$  which is 0.

(2) For  $k$  fixed, it is clear from (1) that  $r_1 \cdots r_n \varepsilon_k = r_1 \cdots r_k \varepsilon_k$  which is  $u\varepsilon_k$  by definition of  $r_k$ .  $\square$

**Proposition 2.7.** *For  $k = 1, \dots, n-1$ , the matrix of the restriction of  $r_k$  to the basis  $(\varepsilon_k, \varepsilon_{k+1})$  is  $\Xi(\gamma_{k-1})$  as defined in (2.14). In particular*

$$\langle \varepsilon_k, r_k \varepsilon_k \rangle = \gamma_{k-1}. \quad (2.17)$$

The restriction of  $r_n$  to  $\mathbb{C}\varepsilon_n$  is the multiplication by  $\gamma_{n-1}$ .

*Proof.* Note that for every  $k \leq n-1$

$$\langle \varepsilon_{k+1}, r_k \varepsilon_k \rangle = \langle r_1 \cdots r_{k-1} \varepsilon_{k+1}, u\varepsilon_k \rangle = \langle \varepsilon_{k+1}, u\varepsilon_k \rangle = \rho_{k-1}. \quad (2.18)$$

Since  $r_k$  is a reflection acting on the subspace  $\text{Span}\{\varepsilon_k, \varepsilon_{k+1}\}$ , identities (2.18) and (2.17) entail that the matrix representing  $r_k$  in the basis  $(\varepsilon_1, \dots, \varepsilon_n)$  is precisely  $\Xi(\gamma_{k-1})$  (see (2.14)). It is then enough to prove (2.17).

For  $k=1$  it is immediate that:

$$\langle \varepsilon_1, r_1 \varepsilon_1 \rangle = \langle \varepsilon_1, u\varepsilon_1 \rangle = \bar{\alpha}_0 = \gamma_0.$$

Let us proceed by induction. For  $j \geq 1$  set  $q_j := \langle \varepsilon_j, r_j \varepsilon_j \rangle$ . Assume that  $q_j = \gamma_{j-1}$  for  $j \leq k$ . We have  $q_{k+1} = \langle \varepsilon_{k+1}, r_{k+1} \varepsilon_{k+1} \rangle = \langle r_1 \cdots r_k \varepsilon_{k+1}, u\varepsilon_{k+1} \rangle$ . Equation (2.12) implies

$$r_k \varepsilon_{k+1} = \varepsilon_{k+1} - \frac{1}{1 - \bar{\gamma}_{k-1}} (r_k \varepsilon_k - \varepsilon_k) \langle r_k \varepsilon_k, \varepsilon_{k+1} \rangle, \quad (2.19)$$

and since  $r_j \varepsilon_\ell = \varepsilon_\ell$  for  $\ell \geq j+2$ , we get:

$$r_1 \cdots r_k \varepsilon_{k+1} = \varepsilon_{k+1} - \frac{1}{1 - \bar{\gamma}_{k-1}} (r_1 \cdots r_k \varepsilon_k - r_1 \cdots r_{k-1} \varepsilon_k) \langle u\varepsilon_k, \varepsilon_{k+1} \rangle.$$

Now, it is known that  $\langle u\varepsilon_k, \varepsilon_{k+1} \rangle = \overline{\langle \varepsilon_{k+1}, u\varepsilon_k \rangle} = \rho_k$ . If we set  $v_1 = \varepsilon_1$ ,

$$v_j = r_1 \cdots r_{j-1} \varepsilon_j, \quad a_j = \frac{\rho_{j-1}}{1 - \bar{\gamma}_{j-1}}, \quad w_{j+1} = \varepsilon_{j+1} - a_j u\varepsilon_j$$

we get the recursion

$$v_{j+1} = a_j v_j + w_{j+1}, \quad (j \leq k), \quad (2.20)$$

which we solve in :

$$v_{k+1} = \left( \prod_{j=1}^k a_j \right) \varepsilon_1 + \sum_{\ell=2}^{k+1} \left( \prod_{j=\ell}^k a_j \right) w_\ell. \quad (2.21)$$

Taking the scalar product with  $u\varepsilon_{k+1}$  yields

$$q_{k+1} = \left( \prod_{j=1}^k \bar{a}_j \right) \langle \varepsilon_1, u\varepsilon_{k+1} \rangle + \sum_{\ell=2}^{k+1} \left( \prod_{j=\ell}^k \bar{a}_j \right) \langle w_\ell, u\varepsilon_{k+1} \rangle.$$

But  $\langle w_\ell, u\varepsilon_{k+1} \rangle = \langle \varepsilon_\ell, u\varepsilon_{k+1} \rangle - \bar{a}_{\ell-1} \langle u\varepsilon_{\ell-1}, u\varepsilon_{k+1} \rangle$ , and since  $\ell \leq k+1$ , we have

$$\langle w_\ell, u\varepsilon_{k+1} \rangle = \langle \varepsilon_\ell, u\varepsilon_{k+1} \rangle = -\bar{\alpha}_k \alpha_{\ell-2} \prod_{m=\ell-1}^{k-1} \rho_m,$$

which yields (with  $\alpha_{-1} = -1$ )

$$\begin{aligned} -\frac{q_{k+1}}{\bar{\alpha}_k} &= \sum_{\ell=1}^{k+1} \left( \prod_{j=\ell}^k \bar{a}_j \right) \alpha_{\ell-2} \prod_{m=\ell-1}^{k-1} \rho_m \\ &= \sum_{\ell=1}^{k+1} \prod_{m=\ell-1}^{k-1} \rho_m^2 \prod_{j=0}^{\ell-3} (1 - \bar{\gamma}_j) \frac{\bar{\gamma}_{\ell-2}}{\prod_{s=0}^{k-1} (1 - \gamma_s)} (1 - \gamma_{\ell-2}) \\ &= \frac{1}{\prod_{s=0}^{k-1} (1 - \gamma_s)} \sum_{\ell=1}^{k+1} \left[ \prod_{m=\ell-2}^{k-1} \rho_m^2 \prod_{j=0}^{\ell-3} (1 - \bar{\gamma}_j) - \prod_{m=\ell-1}^{k-1} \rho_m^2 \prod_{j=0}^{\ell-2} (1 - \bar{\gamma}_j) \right] \\ &= -\prod_{s=0}^{k-1} \frac{(1 - \bar{\gamma}_s)}{(1 - \gamma_s)}, \end{aligned}$$

and eventually  $q_{k+1} = \gamma_k$ .  $\square$

Now, we can summarize the above results in the following theorem.

**Theorem 2.8.** *Let  $u \in U(n)$  and  $e$  a cyclic vector for  $u$ . Let  $\mu$  be the spectral measure of the pair  $(u, e)$ , and  $(\alpha_0, \dots, \alpha_{n-1})$  its Verblunsky coefficients. Let  $(\varepsilon_1, \dots, \varepsilon_n)$  be the orthonormal basis obtained from the Gram-Schmidt procedure applied to  $(e, ue, \dots, u^{n-1}e)$ . Then,  $u$  can be decomposed as a product of  $n$  reflections  $(r_k)_{1 \leq k \leq n}$ :*

$$u = r_1 \dots r_n \tag{2.22}$$

where  $r_1$  is the reflection mapping  $\varepsilon_1$  onto  $u\varepsilon_1$  and by induction for each  $2 \leq k \leq n$ ,  $r_k$  maps  $\varepsilon_k$  onto  $r_{k-1}^{-1} r_{k-2}^{-1} \dots r_1^{-1} u\varepsilon_k$ .

This decomposition can also be restated in terms of the GGT matrix :

$$\mathcal{G}(\alpha_0, \dots, \alpha_{n-1}) = \Xi^{(0)}(\gamma_0) \Xi^{(1)}(\gamma_1) \dots \Xi^{(n-1)}(\gamma_{n-1}), \tag{2.23}$$

where for  $0 \leq k \leq n-2$ , the matrix  $\Xi^{(k)}$  is given by

$$\Xi^{(k-1)}(\gamma_{k-1}) = \text{Id}_{k-1} \oplus \Xi(\gamma_{k-1}) \oplus \text{Id}_{n-k-1}, \tag{2.24}$$

with  $\Xi(\gamma)$  defined in (2.14). For  $k = n-1$ ,

$$\Xi^{(n-1)}(\gamma_{n-1}) = \text{Id}_{n-1} \oplus (\gamma_{n-1}). \tag{2.25}$$

### 3. DEFORMED VERBLUNSKY COEFFICIENTS AND INDEPENDENCE

We now use the point of view of sampling (or change of probability measure) to compute the distribution of the deformed Verblunsky coefficients

under  $\text{CJ}_{\delta,\beta}^{(n)}$ . Let us first remark that, if the  $\alpha_k$ 's are independent and with rotational invariant distribution, then from (2.9)

$$(\alpha_0, \dots, \alpha_{n-1}) \stackrel{\text{law}}{=} (\gamma_0, \dots, \gamma_{n-1}). \quad (3.1)$$

This is the case under  $\text{CJ}_{0,\beta}^{(n)}$ .

We first prove that when  $\delta \neq 0$  the Verblunsky coefficients are not independent anymore by studying the simple case  $n = 2, \beta = 2$ , and then we compute the distribution of  $(\gamma_0, \dots, \gamma_{n-1})$  under  $\text{CJ}_{\delta,\beta}^{(n)}$ . We then show that under this distribution, the weights of the measure associated to the Verblunsky coefficients  $(\alpha_0, \dots, \alpha_{n-1})$  are independent from the points at which the measure is supported and follow a Dirichlet distribution.

Let  $\delta \in \mathbb{C}$  such that  $\Re \delta > -1/2$ . The formula

$$\lambda^{(\delta)}(\zeta) = \frac{\Gamma(1+\delta)\Gamma(1+\bar{\delta})}{\Gamma(1+\delta+\bar{\delta})} (1-\zeta)^{\bar{\delta}} (1-\bar{\zeta})^{\delta}, \quad \zeta \in \mathbb{T} \quad (3.2)$$

defines a probability density with respect to the Haar measure on  $\mathbb{T}$ , which is discontinuous at 1 when  $\Im \delta \neq 0$  (see [8]).

When  $\beta = 2$  and  $\delta \neq 0$ , the Verblunsky coefficients are dependent. Indeed, let  $M \in U(2)$  with Verblunsky coefficients  $\alpha_0$  and  $\alpha_1$ . Then

$$\det(\text{Id} - M) = [1 - \bar{\alpha}_0 - \bar{\alpha}_1(1 - \alpha_0)],$$

with  $|\alpha_0| < 1$  and  $|\alpha_1| = 1$ . Under  $\text{CJ}_{0,2}^{(2)}$ , the variables  $\alpha_0$  and  $\alpha_1$  are independent and uniformly distributed on  $\mathbb{D}$  and  $\mathbb{T}$  respectively (see [24] or Proposition 1.7). The  $\text{CJ}_{\delta,2}^{(2)}$  is a  $\det_{\delta}$  sampling of  $\text{CJ}_{0,2}^{(n)}$  (see the Introduction for the definition and notation for the  $\det_{\delta}$ -sampling). So, the joint density of  $(\alpha, \varphi)$  on  $\mathbb{D} \times \mathbb{T}$  is proportional to

$$\begin{aligned} f(\alpha_0, \alpha_1) &= [1 - \bar{\alpha}_0 - \bar{\alpha}_1(1 - \alpha_0)]^{\bar{\delta}} [1 - \alpha_0 - \alpha_1(1 - \bar{\alpha}_0)]^{\delta} \\ &= (1 - \bar{\alpha}_0)^{\bar{\delta}} (1 - \alpha_0)^{\delta} [1 - \gamma \bar{\alpha}_1]^{\bar{\delta}} [1 - \bar{\gamma} \alpha_1]^{\delta}. \end{aligned}$$

where

$$\gamma = \frac{1 - \alpha}{1 - \bar{\alpha}}.$$

It is then clear that the conditional density of  $\alpha_1$  given  $\alpha_0$  is

$$\alpha_1 \mapsto \lambda^{(\delta)}(\gamma \bar{\alpha}_1), \quad \alpha_1 \in \mathbb{T},$$

and this last quantity does not depend on  $\alpha_0$  (i.e. on  $\gamma$ ) if and only if  $\delta = 0$ . Otherwise the Verblunsky coefficients  $\alpha_0$  and  $\alpha_1$  are dependent.

The next theorem illustrates our interest in the deformed Verblunsky coefficients: under  $\text{CJ}_{\delta,\beta}^{(n)}$ , they are independent. For the proof of this theorem, we shall need the following lemma which will also be useful when we study limit theorems:



**Lemma 3.1.** *Let  $s, t, \ell \in \mathbb{C}$  such that:  $\Re(s + \ell + 1) > 0, \Re(t + \ell + 1) > 0$ . Then, the following identity holds:*

$$\int_{\mathbb{D}} (1 - |z|^2)^{\ell-1} (1 - z)^s (1 - \bar{z})^t d^2 z = \frac{\pi \Gamma(\ell) \Gamma(\ell + 1 + s + t)}{\Gamma(\ell + 1 + s) \Gamma(\ell + 1 + t)}. \quad (3.3)$$

*Proof.* A Taylor expansion yields

$$(1 - z)^s (1 - \bar{z})^t = \sum_{m, n \geq 0} \rho^{m+n} \frac{(-s)_n (-t)_m}{n! m!} e^{i(m-n)\theta},$$

with  $z = \rho e^{i\theta}$  and  $0 \leq \rho < 1$ . We obtain by integrating

$$\begin{aligned} \int_{\mathbb{D}} (1 - |z|^2)^{\ell-1} (1 - z)^s (1 - \bar{z})^t d^2 z &= 2\pi \sum_{n \geq 0} \frac{(-s)_n (-t)_n}{n! n!} \int_0^1 (1 - \rho^2)^{\ell-1} \rho^{2n+1} d\rho \\ &= \pi \sum_{n \geq 0} \frac{(-s)_n (-t)_n}{n!} \frac{(\ell - 1)!}{(n + \ell)!} = \frac{\pi}{\ell} {}_2F_1(-s, -t; \ell + 1; 1), \end{aligned}$$

where  ${}_2F_1$  is the classical hypergeometric function (see [4]) and an application of Gauss formula (see [4]) shows that the last expression is exactly the right hand side of (3.3).  $\square$

**Theorem 3.2.** *Let  $\delta \in \mathbb{C}$  with  $\Re \delta > -1/2$  and  $\beta > 0$ . Set  $\beta' = \beta/2$ . Under  $CJ_{\delta, \beta}^{(n)}$ , the distribution of  $(\gamma_0, \dots, \gamma_{n-1})$ , denoted hereafter  $\eta_{\delta, \beta}^{(n)}$ , is the following:*

- (1) *the variables  $\gamma_0, \dots, \gamma_{n-2}, \gamma_{n-1}$  are independent ;*
- (2) *for  $k = 0, \dots, n - 2$  the density of  $\gamma_k$  with respect to the Lebesgue measure  $d^2 z$  on  $\mathbb{C}$  is*

$$c_{k,n}(\delta) (1 - |z|^2)^{\beta'(n-k-1)-1} (1 - z)^{\bar{\delta}} (1 - \bar{z})^{\delta} \mathbb{1}_{\mathbb{D}}(z),$$

where

$$c_{k,n}(\delta) = \frac{\Gamma(\beta'(n - k - 1) + 1 + \delta) \Gamma(\beta'(n - k - 1) + 1 + \bar{\delta})}{\pi \Gamma(\beta'(n - k - 1)) \Gamma(\beta'(n - k - 1) + 1 + \delta + \bar{\delta})}; \quad (3.4)$$

- (3) *the density of  $\gamma_{n-1}$  with respect to the Haar measure on  $\mathbb{T}$  is  $\lambda^{(\delta)}$  (given by (3.2)).*

*Proof.* The distribution of the  $\alpha$ 's in the  $\beta$ -circular unitary ensemble is  $\eta_{0, \beta}^{(n)}$ . More precisely, as seen in Definition 1.5 they are independent and if

$$\alpha_k = R_k e^{i\psi_k} \quad (0 \leq k \leq n - 2), \quad \alpha_{n-1} = e^{i\psi_{n-1}},$$

then  $R_k$  and  $\psi_k$  are independent,  $\psi_k$  is uniformly distributed and  $R_k^2$  has the Beta( $1, \beta'(n - k - 1)$ ) distribution. Moreover  $\alpha_{n-1}$  is uniformly distributed on  $\mathbb{T}$ . From (2.11), the sampling factor is

$$\det(\text{Id} - U)^{\bar{\delta}} \det(\text{Id} - \bar{U})^{\delta} = (1 - \gamma_{n-1})^{\bar{\delta}} (1 - \bar{\gamma}_{n-1})^{\delta} \prod_{k=0}^{n-2} (1 - \gamma_k)^{\bar{\delta}} (1 - \bar{\gamma}_k)^{\delta}.$$

so that, under  $\text{CJ}_{\delta,\beta}^{(n)}$ , the density of  $(R_0, \dots, R_{n-2}, \psi_0, \dots, \psi_{n-1})$  is proportional to

$$\lambda^{(\delta)}(\gamma_{n-1}) \prod_{k=0}^{n-2} (1 - R_k^2)^{\beta'(n-1-k)-1} R_k (1 - \gamma_k)^{\bar{\delta}} (1 - \bar{\gamma}_k)^{\delta} \mathbb{1}_{(0,1)}(R_k),$$

with

$$\gamma_k = R_k e^{i\theta_k} \quad (0 \leq k \leq n-2), \quad \gamma_{n-1} = e^{i\theta_{n-1}}.$$

Thanks to the relations (2.9) and (2.10), the Jacobian matrix of the mapping

$$(R_0, \dots, R_{n-2}, \psi_0, \dots, \psi_{n-1}) \rightarrow (R_0, \dots, R_{n-2}, \theta_0, \dots, \theta_{n-1})$$

is lower triangular with diagonal elements  $\pm 1$ , so that, under  $\text{CJ}_{\delta,\beta}^{(n)}$ , the density of  $(R_0, \dots, R_{n-2}, \theta_0, \dots, \theta_{n-1})$ , is proportional to

$$\lambda^{(\delta)}(\gamma_{n-1}) \prod_{k=0}^{n-2} (1 - R_k^2)^{\beta'(n-1-k)-1} R_k (1 - \gamma_k)^{\bar{\delta}} (1 - \bar{\gamma}_k)^{\delta} \mathbb{1}_{(0,1)}(R_k), \quad (3.5)$$

which proves the required independence and the expression of the distributions, up to the determination of  $c_{k,n}(\delta)$ . This quantity is obtained by taking  $\ell = \beta'(n-k-1)$ ,  $s = \bar{\delta}$ ,  $t = \delta$  in (3.3), which gives (3.4) and completes the proof of the Theorem.  $\square$

Starting with a set of deformed Verblunsky coefficients, with distribution  $\eta_{\delta,\beta}^{(n)}$ , we obtain the coefficients  $(\alpha_0, \dots, \alpha_{n-1})$  by inverting formula (2.11). These are the coordinates of some probability measure  $\mu$  supported at  $n$  points on the unit circle:

$$\mu = \sum_{k=1}^n \pi_k \delta_{e^{i\theta_k}},$$

with  $\pi_k > 0$  and  $\sum_{k=1}^n \pi_k = 1$ . The next theorem gives the distribution induced on the vector  $(\pi_1, \dots, \pi_n, \theta_1, \dots, \theta_n)$  by  $(\gamma_0, \dots, \gamma_{n-1})$ .

**Theorem 3.3.** *The following formulae express the same measure on the manifold of probability distribution on  $\mathbb{T}$  supported at  $n$  points:*

$$K_{\delta,\beta}^{(n)} |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta \prod_{k=1}^n (1 - e^{-i\theta_k})^\delta (1 - e^{i\theta_k})^{\bar{\delta}} \prod_{k=1}^n \pi_k^{\beta'-1} d\theta_1 \dots d\theta_n d\pi_1 \dots d\pi_{n-1}$$

in the  $(\theta, \pi)$  coordinates and

$$K_{\delta,\beta}^{(n)} \prod_{k=0}^{n-2} (1 - |\gamma_k|^2)^{\beta'(n-k-1)-1} \prod_{k=0}^{n-1} (1 - \gamma_k)^{\bar{\delta}} (1 - \bar{\gamma}_k)^{\delta} d^2 \gamma_0 \dots d^2 \gamma_{n-2} d\phi$$

in terms of the deformed Verblunsky coefficients, with  $\gamma_{n-1} = e^{i\phi}$ . Here,  $K_{\delta,\beta}^{(n)}$  is a constant:

$$K_{\delta,\beta}^{(n)} = \frac{\Gamma(1+\delta)\Gamma(1+\bar{\delta})}{2^{n-1}\pi\Gamma(1+\delta+\bar{\delta})} \prod_{k=0}^{n-2} c_{k,n}(\delta),$$

with  $c_{k,n}(\delta)$  given in Theorem 3.2. Consequently, if  $(\gamma_0, \dots, \gamma_{n-1})$  is  $\eta_{\delta,\beta}^{(n)}$  distributed, then  $(\pi_1, \dots, \pi_n)$  and  $(\theta_1, \dots, \theta_n)$  are independent; the vector of weights  $(\pi_1, \dots, \pi_n)$  follows the  $\text{Dir}_n(\beta')$  distribution and the vector  $(\theta_1, \dots, \theta_n)$  has the density  $h_{\delta,\beta}^{(n)}$ .

*Proof.* In the course of this proof, we shall adopt the following point of view. Starting with a measure supported at  $n$  points on the unit circle, we associate with it its Verblunsky coefficients  $(\alpha_0, \dots, \alpha_{n-1})$  and then the corresponding GGT matrix which we note  $G$  for simplicity. Then  $e_1$  is a cyclic vector for  $G$  and  $\mu$  is the spectral measure of  $(G, e_1)$ . Conversely, starting with the set of deformed Verblunsky coefficients with  $\eta_{\delta,\beta}$  distribution, we construct the coefficients  $(\alpha_0, \dots, \alpha_{n-1})$  with the transformations (2.9), then the GGT matrix associated with it and finally  $\mu$  the spectral measure associated with this matrix and  $e_1$ .

We use the following well-known identity (see [35] or [24] Lemma 4.1):

$$|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^2 \prod_{k=1}^n \pi_k = \prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{n-k-1}. \quad (3.6)$$

Since  $|\gamma_k| = |\alpha_k|$ , we can also write

$$|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^2 \prod_{k=1}^n \pi_k = \prod_{k=0}^{n-2} (1 - |\gamma_k|^2)^{n-k-1}. \quad (3.7)$$

Moreover, from (2.11),

$$\det(\text{Id} - G) = \prod_{k=1}^n (1 - e^{i\theta_k}) = \prod_{k=0}^{n-1} (1 - \gamma_k). \quad (3.8)$$

In our setting,  $\pi_i$  is modulus squared of the first component of the  $i$ -th eigenvector of the matrix  $G$ . Now, define

$$q_k^2 = \pi_k, \quad \text{for } k = 1, \dots, n.$$

It is known (see for example Forrester [13], Chapter 2 and [14] Theorem 2) that the Jacobian of the map  $(\alpha_0, \dots, \alpha_{n-1}) \mapsto (\theta_1, \dots, \theta_n, q_1, \dots, q_{n-1})$  is given by

$$\frac{\prod_{k=0}^{n-2} (1 - |\alpha_k|^2)}{q_n \prod_{k=1}^n q_k}.$$

Moreover, the map  $(\gamma_0, \dots, \gamma_{n-1}) \mapsto (\alpha_0, \dots, \alpha_{n-1})$  is invertible and its Jacobian is 1, as already seen. The result now follows from simple integral manipulations combined with the identities (3.7) and (3.8).  $\square$

#### 4. MATRIX MODELS FOR THE CIRCULAR JACOBI ENSEMBLE

The results of the previous sections can now be used to propose some simple matrix models for the circular Jacobi ensemble. There are mainly two ways to generate matrix models for a given spectral measure encoded by its Verblunsky coefficients.

*The AGR decomposition* : if  $U = \Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\dots\Theta^{(n-1)}(\alpha_{n-1})$  (the  $\Theta_k$ 's are defined in the introduction), the Verblunsky coefficients for the spectral measure associated to  $(U, e_1)$  are precisely  $(\alpha_0, \dots, \alpha_{n-1})$  (see [3] or [33] Section 10). Therefore, taking independent  $\alpha_k$ 's with law  $\eta_{0,\beta}^{(n)}$ , the density of the eigenvalues of

$$U = \Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\dots\Theta^{(n-1)}(\alpha_{n-1})$$

is proportional to  $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta$ . The matrix  $U$  obtained above is the GGT matrix associated with the  $\alpha_k$ 's. It is in the Hessenberg form.

*The CMV form* : set

$$\begin{cases} \mathcal{L} &= \Theta^{(0)}(\alpha_0)\Theta^{(2)}(\alpha_2)\dots \\ \mathcal{M} &= \Theta^{(1)}(\alpha_1)\Theta^{(3)}(\alpha_3)\dots \end{cases}$$

Cantero, Moral, and Velazquez [9] proved that the Verblunsky coefficients associated to  $(\mathcal{LM}, e_1)$  are precisely  $(\alpha_0, \dots, \alpha_{n-1})$ . Therefore, taking as previously independent  $\alpha_k$ 's with distribution  $\eta_{0,\beta}^{(n)}$ , the density of the eigenvalues of the spectral law of  $\mathcal{LM}$  is proportional to  $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^\beta$  ([24]). This matrix model is very sparse: it is pentadiagonal.

We now propose a matrix model for the circular Jacobi ensemble: it is reminiscent of the AGR factorization with the noticeable difference that it is based on the deformed Verblunsky coefficients and actual reflections as defined in Section 2.

**Theorem 4.1.** *If  $(\gamma_0, \dots, \gamma_{n-1})$  is  $\eta_{\delta,\beta}^{(n)}$  distributed, then with the notation of (2.24) and (2.25),*

$$\Xi^{(0)}(\gamma_0)\Xi^{(1)}(\gamma_1)\dots\Xi^{(n-1)}(\gamma_{n-1})$$

*is a matrix model for the circular Jacobi ensemble, i.e. the density of the eigenvalues is  $h_{\delta,\beta}^{(n)}$  (see (1.6)).*

*Proof.* We know from Theorem 2.8 that

$$\mathcal{G}(\alpha_0, \dots, \alpha_{n-1}) = \Xi^{(0)}(\gamma_0)\Xi^{(1)}(\gamma_1)\dots\Xi^{(n-1)}(\gamma_{n-1}). \quad (4.1)$$

We also proved in Theorem 3.3 that the set of deformed Verblunsky coefficients with probability distribution  $\eta_{\delta,\beta}^{(n)}$  induces a distribution on the eigenvalues of the GGT matrix  $\mathcal{G}(\alpha_0, \dots, \alpha_{n-1})$  which has exactly the density  $h_{\delta,\beta}$ . This completes the proof of the Theorem.  $\square$

*Remark.* We now say a few extra words on the CMV form obtained by Killip and Nenciu in [24]. Cantero, Moral and Velazquez [9] introduced the basis  $\chi_0, \dots, \chi_{n-1}$  obtained by orthogonalizing the sequence  $1, z, z^{-1}, \dots$ . They prove that in this basis the matrix is pentadiagonal. We name this matrix  $\mathcal{C}(\alpha_0, \dots, \alpha_{n-1})$ . It turns out that there exists a unitary  $P$  such that:

$$P\mathcal{G}(\alpha_0, \dots, \alpha_{n-1})P^* = \mathcal{C}(\alpha_0, \dots, \alpha_{n-1}), \quad P\varphi_0 = \chi_0.$$

The two pairs  $(\mathcal{G}(\alpha_0, \dots, \alpha_{n-1}), \varphi_0)$  and  $(\mathcal{C}(\alpha_0, \dots, \alpha_{n-1}), \chi_0)$  are equivalent, they admit the  $\alpha_k$ 's as Verblunsky coefficients, and have the same spectral measure. We conclude that if we start with the  $\gamma_k$ 's distributed as  $\eta_{\delta, \beta}^{(n)}$ , and build the  $\alpha_k$ 's by inverting the transformation (2.9), then  $\mathcal{C}(\alpha_0, \dots, \alpha_{n-1})$  will be a matrix model for the circular Jacobi ensemble. But we do not know how to construct the CMV matrix from the  $\gamma_k$ 's directly. We saw at the beginning of this section that Cantero et al. introduced the matrices  $\mathcal{L}$  and  $\mathcal{M}$ , as direct product of small blocks  $\Theta^{(k)}(\alpha_k)$  and obtained  $\mathcal{C}$  as  $\mathcal{C} = \mathcal{L}\mathcal{M}$ . It would be interesting to have an analogue construction based on the independent  $\gamma_k$ 's.

Theorem 2.8 which is a deterministic result, has also the following consequence:

**Proposition 4.2.** *Let  $(\alpha_0, \dots, \alpha_{n-2}, \alpha_{n-1}) \in \mathbb{D}^{n-1} \times \mathbb{T}$  be independent random variables with rotationally invariant distribution. Then*

$$\Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\dots\Theta^{(n-1)}(\alpha_{n-1}) \stackrel{\text{law}}{=} \Xi^{(0)}(\alpha_0)\Xi^{(1)}(\alpha_1)\dots\Xi^{(n-1)}(\alpha_{n-1}).$$

*Proof.* We give two proofs of this result. The first one is a consequence of Theorem 2.8 from which we know that

$$\Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\dots\Theta^{(n-1)}(\alpha_{n-1}) = \Xi^{(0)}(\gamma_0)\Xi^{(1)}(\gamma_1)\dots\Xi^{(n-1)}(\gamma_{n-1}),$$

and the remark at the beginning of Section 3.

For the second proof, we proceed by induction on  $n$ . For  $n = 1$  the result is obvious. Suppose the result holds at rank  $n - 1$  : thanks to the recurrence hypothesis,

$$\Xi^{(0)}(\alpha_0)\Xi^{(1)}(\alpha_1)\dots\Xi^{(n-1)}(\alpha_{n-1}) \stackrel{\text{law}}{=} \Xi^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\dots\Theta^{(n-1)}(\alpha_{n-1}).$$

Let  $e^{i\phi_0} = \frac{1-\bar{\alpha}_0}{1-\alpha_0}$ . An elementary calculation gives

$$\begin{aligned} & \Xi^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\dots\Theta^{(n-2)}(\alpha_{n-2})\Theta^{(n-1)}(\alpha_{n-1}) \\ &= \Theta^{(0)}(\alpha_0)\Theta^{(1)}(e^{-i\phi_0}\alpha_1)\dots\Theta^{(n-2)}(e^{-i\phi_0}\alpha_{n-1})\Theta^{(n-1)}(e^{-i\phi_0}\alpha_{n-1}). \end{aligned}$$

As the  $\alpha_k$ 's are independent with law invariant by rotation,

$$(\alpha_0, e^{-i\phi_0}\alpha_1, \dots, e^{-i\phi_0}\alpha_{n-2}, e^{i\phi_0}\alpha_{n-1}) \stackrel{\text{law}}{=} (\alpha_0, \alpha_1, \dots, \alpha_{n-2}, \alpha_{n-1}),$$

which completes the proof. □

Now that we have a matrix model for the circular Jacobi ensemble, we can study the characteristic polynomial for such matrices; the key formula will be (2.11).

**Proposition 4.3.** *Let  $U$  be a unitary matrix of size  $n$  and let  $Z_n = \det(\text{Id} - U)$  be its characteristic polynomial evaluated at 1. Then, in the circular*

Jacobi ensemble,  $Z_n$  can be written as a product of  $n$  independent complex random variables:

$$Z_n = \prod_{k=0}^{n-1} (1 - \gamma_k),$$

where the laws of the  $\gamma_k$ 's are given in Theorem 3.2. Consequently for any  $s, t \in \mathbb{C}$ , with  $\Re(t) > -\frac{1}{2}$ , the Mellin-Fourier transform of  $Z_n$  is :

$$\mathbb{E}[|Z_n|^t e^{is \arg Z_n}] = \prod_{k=0}^{n-1} \frac{\Gamma(\beta'k + 1 + \delta)\Gamma(\beta'k + 1 + \bar{\delta})\Gamma(\beta'k + 1 + \delta + \bar{\delta} + t)}{\Gamma(\beta'k + 1 + \delta + \bar{\delta})\Gamma(\beta'k + 1 + \delta + \frac{t-s}{2})\Gamma(\beta'k + 1 + \bar{\delta} + \frac{t+s}{2})}. \quad (4.2)$$

*Proof.* The first part is an easy consequence of (2.11) and Theorem 3.2. To prove the second part, we note that if  $X_k = (1 - \gamma_k)$ , then  $|X_k|^t e^{is \arg X_k} = (1 - \gamma_k)^a (1 - \bar{\gamma}_k)^b$ , where  $a = (t + s)/2$  and  $b = (t - s)/2$ . Consequently, by independence of the  $\gamma_k$ 's, we obtain:

$$\mathbb{E}[|Z_n|^t e^{is \arg Z_n}] = \prod_{k=0}^{n-1} \mathbb{E}[(1 - \gamma_k)^a (1 - \bar{\gamma}_k)^b],$$

and formula (4.2) then easily follows from Lemma 3.1.  $\square$

## 5. LIMITING SPECTRAL MEASURE AND LARGE DEVIATIONS

In (1.7) we defined the spectral measure which is a central tool for the study of our circular ensembles. Let us re-write this measure on  $\mathbb{T}$  as

$$\mu_{\text{sp}}^{(n)} := \sum_{k=1}^n \pi_k^{(n)} \delta_{e^{i\theta_k^{(n)}}}, \quad (5.1)$$

where we put a superscript  $(n)$  to stress on the dependency on  $n$ . Besides, in classical Random Matrix Theory, many authors are mainly interested in the empirical spectral distribution (ESD) defined by

$$\mu_{\text{esd}}^{(n)} = \frac{1}{n} \sum_{k=1}^n \delta_{e^{i\theta_k^{(n)}}}. \quad (5.2)$$

We are concerned with their asymptotics under  $\text{CJ}_{\delta, \beta}^{(n)}$  when  $n \rightarrow \infty$  with

$$\delta = \delta(n) = \beta' nd$$

where  $\Re d \geq 0$  (and as usual  $\beta' = \beta/2$ ). In this framework the variables  $(\theta_1^{(n)}, \dots, \theta_n^{(n)})$  and  $(\pi_1^{(n)}, \dots, \pi_n^{(n)})$  are distributed as in Theorem 3.3. It is well-known that in the CUE (i.e. for  $d = 0$  and  $\beta' = 1$ ), the sequence  $(\mu_{\text{esd}}^{(n)})$  converges weakly in probability to the uniform measure on  $\mathbb{T}$  and that it satisfies the Large Deviation Principle at scale  $n^2$  ([17] chap. 5).

In this section we prove that in the cJUE, both sequences  $(\mu_{\text{sp}}^{(n)})$  and  $(\mu_{\text{esd}}^{(n)})$  converge weakly in probability to the same limit measure supported by an

arc of the unit circle, and that  $(\mu_{\text{esd}}^{(n)})$  satisfies the Large Deviation Principle at scale  $n^2$ . In the next two subsections, we first recall some definitions (which may be skipped by probabilist readers) and then state the main results which are Theorems 5.1, 5.4 and 5.5.

**5.1. Weak convergence of measures.** Let  $\mathcal{M}_1(\mathbb{T})$  be the space of probability measures on  $\mathbb{T}$ . The weak topology on  $\mathcal{M}_1(\mathbb{T})$  is defined from the duality with the space of continuous functions on  $\mathbb{T}$ ; that is,  $\nu_n$  weakly converges to  $\nu$  means  $\int f d\nu_n \rightarrow \int f d\nu$  for every continuous function  $f$ . Since  $\mathbb{T}$  is compact, it is equivalent to the convergence of moments. If we define the distribution function of  $\nu \in \mathcal{M}_1(\mathbb{T})$  as  $F_\nu(t) = \nu(\{e^{i\theta}; \theta \in [0, t]\})$ , then  $\nu_n \rightarrow \nu$  weakly if and only if

$$F_{\nu_n}(t) \rightarrow F_\nu(t)$$

for all  $t \in [0, 2\pi)$  at which  $F_\nu(t)$  is continuous. The Lévy distance<sup>2</sup>  $d_L$  is compatible with this topology and makes  $\mathcal{M}_1(\mathbb{T})$  a compact metric space. It should be noticed that for all pair of elements of  $\mathcal{M}_1(\mathbb{T})$

$$d_L(\mu, \nu) \leq \sup_t |F_\mu(t) - F_\nu(t)|. \quad (5.3)$$

In the following we consider random probability measures  $(\nu_n)$  and the weak convergence in probability of  $(\nu_n)$  to a deterministic  $\nu \in \mathcal{M}_1(\mathbb{T})$ . That means that for every  $n \geq 1$  we have a probability space  $(\Omega_n, \mathcal{F}_n, Q_n)$ , a  $\mathcal{M}_1(\mathbb{T})$ -valued random variable  $\nu_n$ , and that  $d(\nu_n, \nu) \rightarrow 0$  in probability, where  $d$  is any distance compatible with the weak topology (e.g. the Lévy distance  $d_L$ ). In other words

$$\begin{aligned} \lim_n \nu_n = \nu \quad (\text{in probability}) &\Leftrightarrow \lim_n d(\nu_n, \nu) = 0 \quad (\text{in probability}) \\ &\Leftrightarrow \forall \varepsilon > 0 \quad \lim_n Q_n(d(\nu_n, \nu) > \varepsilon) = 0. \end{aligned} \quad (5.4)$$

**5.1.1. The spectral measure.** The following theorem states the convergence of the sequence  $(\mu_{\text{sp}}^{(n)})$  to an explicit limit, which we identify now, with the help of some more notation. We assume that  $\Re d \geq 0$ . Let

$$\alpha_d = -\frac{\bar{d}}{1 + \bar{d}} \quad (5.5)$$

and let  $\theta_d \in [0, \pi)$  and  $\xi_d \in [-|\theta_d|, |\theta_d|]$  be such that

$$\sin \frac{\theta_d}{2} = \left| \frac{d}{1 + d} \right|, \quad e^{i\xi_d} = \frac{1 + d}{1 + \bar{d}}.$$

---

<sup>2</sup>whose precise definition is not needed here (see [10] Theorem D8)

If  $\Re d \geq 0, d \neq 0$ , let  $w_d$  be defined by

$$w_d(\theta) = \begin{cases} \frac{\sqrt{\sin^2((\theta - \xi_d)/2) - \sin^2(\theta_d/2)}}{|1 + \alpha_d| \sin(\theta/2)} & \text{if } \theta \in (\theta_d + \xi_d, 2\pi - \theta_d + \xi_d) \\ 0 & \text{otherwise.} \end{cases} \quad (5.6)$$

In this case let

$$d\mu_d(\zeta) := w_d(\theta) \frac{d\theta}{2\pi} \quad (\zeta = e^{i\theta}), \quad (5.7)$$

and for  $d = 0$  let

$$d\mu_0(\zeta) := \frac{d\theta}{2\pi} \quad (\zeta = e^{i\theta}),$$

the Haar measure on  $\mathbb{T}$ .

**Theorem 5.1.** *Assume  $\Re d \geq 0$ . As  $n \rightarrow \infty$ ,*

$$\lim \mu_{\text{sp}}^{(n)} = \mu_d \quad (\text{in probability}).$$

To prove this convergence we use the parameterization of measures by their modified Verblunsky coefficients and the following two lemmas whose proofs are postponed until the end of the section.

**Lemma 5.2.** *For every fixed  $k \geq 0$ , as  $n \rightarrow \infty$ ,*

$$\lim \gamma_k^{(n)} = -\frac{d}{1 + \bar{d}} \quad (\text{in probability}), \quad (5.8)$$

and consequently

$$\lim \alpha_k^{(n)} = \alpha_d e^{-i(k+1)\xi_d} \quad (\text{in probability}). \quad (5.9)$$

**Lemma 5.3.** *The sequence of Verblunsky coefficients of  $\mu_d$  is precisely  $(\alpha_d e^{-i(k+1)\xi_d})_{k \geq 0}$ .*

*Proof of Theorem 5.1.* For  $\nu \in \mathcal{M}_1(\mathbb{T})$ , let  $m_k(\nu) = \int e^{ik\theta} d\nu(\theta)$  be its  $k$ -th moment and let  $\alpha_k(\nu)$  be its  $k$ -th Verblunsky - in short V- coefficient. Moments are related to V-coefficients in a continuous way : for every  $j \geq 1$ ,  $m_j(\nu)$  is a continuous function of  $(\alpha_0, \dots, \alpha_{(j \wedge N)-1})$  where  $N$  is the cardinal of the support of  $\nu$ . We know for Lemma 5.2 that for every fixed  $k$ ,  $\lim \alpha_j(\mu_{\text{sp}}^{(n)}) = \alpha_j(\mu_d)$  for  $j \leq k$ , in probability. It entails the convergence of  $m_k(\mu_{\text{sp}}^{(n)})$  to  $m_k(\mu_d)$ , in probability, hence the weak convergence of the measures.  $\square$

*Remark.* The above method can easily be adapted to show the trivial asymptotics of two different scaling regimes :

- if  $\delta(n) = o(n)$ , the limit is the uniform measure on  $\mathbb{T}$ ,
- if  $\delta(n)$  is real and  $n = o(\delta(n))$ , the limit is the Dirac measure at  $-1$ .



*Proof of Lemma 5.2.* For  $\gamma_k^{(n)}$  we use the Mellin transform

$$\mathbb{E} \left( (1 - \gamma_k^{(n)})^s \right) = \frac{\Gamma(\beta'(n - k - 1) + \delta + \bar{\delta} + s + 1) \Gamma(\beta'(n - k - 1) + \bar{\delta} + 1)}{\Gamma(\beta'(n - k - 1) + \delta + \bar{\delta} + 1) \Gamma(\beta'(n - k - 1) + \bar{\delta} + s + 1)},$$

(this comes immediately from (3.3)).

Since for fixed  $z \in \mathbb{C}$

$$\lim_{n \rightarrow \infty} \frac{\Gamma(n + z)}{\Gamma(n) n^z} = 1,$$

we get, for fixed  $s$  (and  $k$ )

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( (1 - \gamma_k^{(n)})^s \right) = \left( \frac{1 + \bar{d} + d}{1 + \bar{d}} \right)^s,$$

which implies that

$$\lim_{n \rightarrow \infty} (1 - \gamma_k^{(n)}) = \frac{1 + \bar{d} + d}{1 + \bar{d}} \text{ in probability,}$$

which is equivalent to (5.8). The statement (5.9) is a direct consequence of (5.8) and (2.9).  $\square$

*Remark.* The convergences in probability in the above lemma actually hold in  $L^p$ , for all  $p > 0$  because all variables are bounded by 1.

*Proof of Lemma 5.3.* If  $d$  is real, the Verblunsky coefficients are all equal. In fact, measures satisfying this property are known to have an absolute continuous part supported by an arc of the unit circle, and a possible additional Dirac mass ([35] p.87). More precisely, let  $\alpha \in \mathbb{D}$  and let  $\nu$  be the measure on  $\mathbb{T}$  such that  $\alpha_k(\nu) = \alpha$  for every  $k \geq 0$ . If  $\theta(\alpha)$  and  $\xi(\alpha)$  are defined by

$$\theta(\alpha) = 2 \arcsin |\alpha|, \quad e^{i\xi(\alpha)} = \frac{1 + \alpha}{1 + \bar{\alpha}},$$

then

$$d\nu(\zeta) = w(\theta) \frac{d\theta}{2\pi} \quad (\zeta = e^{i\theta}),$$

where

$$w(\theta) = \begin{cases} \frac{\sqrt{\sin^2(\theta/2) - \sin^2(\theta(\alpha)/2)}}{|1 + \alpha| \sin((\theta + \xi(\alpha))/2)} & \text{if } \theta \in (\theta(\alpha), 2\pi - \theta(\alpha)), \\ 0 & \text{otherwise} \end{cases} \quad (5.10)$$

as soon as

$$\left| \alpha + \frac{1}{2} \right| \leq \frac{1}{2} \quad (5.11)$$

(otherwise, there is an additional Dirac mass). The orthogonal polynomials with respect to this measure are known as the "Geronimus polynomials".

When  $\alpha = \alpha_d$ , we set  $\theta(\alpha) = \theta_d$ ,  $\xi(\alpha) = \xi_d$  and  $\nu = \nu_d$ ,  $w = w_d$ . We see that

$$\alpha_d + \frac{1}{2} = \frac{1 - \bar{d}}{2(1 + \bar{d})},$$

so that the condition (5.11) is fulfilled if and only if  $\Re d \geq 0$ , which we assumed.

Moreover, it is known (see [36] p.960) that if  $(\alpha_k)_{k \geq 0}$  is the sequence of Verblunsky coefficients of a measure  $\mu$ , then the coefficients  $(e^{-i(k+1)\xi_d} \alpha_k)_{k \geq 0}$  are associated with  $\mu$  rotated by  $\xi_d$ . Consequently,

$$d\mu_d(\zeta) = d\nu_d(\zeta e^{-i\xi_d}),$$

which is precisely (5.7).  $\square$

5.1.2. *The ESD.* In matrix models, the convergence of the ESD is often tackled directly via the convergence of moments or the convergence of the Cauchy transform. Here, we follow a different way: we use Theorem 5.1 and prove that the two sequences  $(\mu_{\text{sp}}^{(n)})_n$  and  $(\mu_{\text{esd}}^{(n)})_n$  are "contiguous".

**Theorem 5.4.** *Assume  $\Re d \geq 0$ . As  $n \rightarrow \infty$ ,*

$$\lim_n \mu_{\text{esd}}^{(n)} = \mu_d \quad (\text{in probability}),$$

where  $\mu_d$  is given in (5.7).

*Proof.* From Theorem 5.1, we know, with the notation from the beginning of Section 5.1 that

$$\lim_{n \rightarrow \infty} d_L(\mu_{\text{sp}}^{(n)}, \mu_d) = 0 \quad \text{in probability.}$$

By the triangle inequality

$$d_L(\mu_{\text{esd}}^{(n)}, \mu_d) \leq d_L(\mu_{\text{sp}}^{(n)}, \mu_d) + d_L(\mu_{\text{sp}}^{(n)}, \mu_{\text{esd}}^{(n)})$$

so, it is enough to prove that  $d_L(\mu_{\text{sp}}^{(n)}, \mu_{\text{esd}}^{(n)})$  converges to 0 in probability. Thanks to (5.3), it is enough to prove

$$\lim_{n \rightarrow \infty} \sup_t |F_{\mu_{\text{sp}}^{(n)}}(t) - F_{\mu_{\text{esd}}^{(n)}}(t)| = 0 \quad \text{in probability.} \quad (5.12)$$

In fact,

$$\sup_t |F_{\mu_{\text{sp}}^{(n)}}(t) - F_{\mu_{\text{esd}}^{(n)}}(t)| = \max_k |S_k^{(n)} - \frac{k}{n}|, \quad (5.13)$$

where  $S_k^{(n)} = \sum_{j=1}^k \pi_j^{(n)}$ . Using the union bound and the Markov inequality, we may write

$$\begin{aligned} \mathbb{P} \left( \max_k \left| S_k^{(n)} - \frac{k}{n} \right| > \varepsilon \right) &\leq \sum_{k=1}^n \mathbb{P} \left( \left| S_k^{(n)} - \frac{k}{n} \right| > \varepsilon \right) \\ &\leq \varepsilon^{-4} \sum_{k=1}^n \mathbb{E} \left( \left( S_k^{(n)} - \frac{k}{n} \right)^4 \right), \end{aligned} \quad (5.14)$$

and then use explicit distributions. Indeed, we showed in Theorem 3.3 that the vector  $(\pi_1^{(n)}, \dots, \pi_n^{(n)})$  follows the  $\text{Dir}_n(\beta')$  distribution. It entails that for  $k = 1, \dots, n-1$ , the variable  $S_k^{(n)}$  is  $\text{Beta}(\beta'k, \beta'(n-k))$  distributed.

Recall that the Mellin transform of a beta variable  $\text{Beta}(a, b)$  with positive parameters  $a$  and  $b$  is

$$\mathbb{E}(\text{Beta}(a, b)^s) = \frac{\Gamma(a+s)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+s)},$$

for  $s > -a$ . As a consequence,  $\mathbb{E} \text{Beta}(a, b) = a/(a+b)$  and a straightforward calculation gives

$$\begin{aligned} \mathbb{E} \left( (\text{Beta}(a, b) - \mathbb{E} \text{Beta}(a, b))^4 \right) &= \frac{3ab(2a^2 + 2b^2 - 2ab + a^2b + ab^2)}{(a+b)^4(a+b+1)(a+b+2)(a+b+3)} \\ &= O\left(\frac{ab}{(a+b)^4}\right). \end{aligned}$$

When  $a = \beta'k$  and  $b = \beta'(n-k)$ , this shows that the  $k$ -th term in the sum of (5.14) is  $O\left(\frac{k(n-k)}{n^4}\right)$ , so that

$$\mathbb{P} \left( \max_k \left| S_k^{(n)} - \frac{k}{n} \right| > \varepsilon \right) = O\left(\sum_{k=1}^n \frac{k(n-k)}{n^4}\right) = O\left(\frac{1}{n}\right).$$

Thanks to (5.13), this yields (5.12) and completes the proof.  $\square$

**5.2. Large deviations for the ESD.** It turns out that the convergence in Theorem 5.4 is exponentially fast, and to make this statement more precise we need to recall the definition of some notions of large deviations, for which the reference is the book of Dembo and Zeitouni ([10]). We say that a sequence  $(P_n)$  of probability measures on a measurable Hausdorff space  $(\mathcal{X}, B(\mathcal{X}))$  satisfies the LDP at scale  $u_n$  (with  $u_n \rightarrow \infty$ ), if there exists a lower semicontinuous function  $I : \mathcal{X} \rightarrow [0, \infty]$  such that

$$\limsup \frac{1}{u_n} \log P_n(F) \leq -\inf\{I(x); x \in F\} \quad (5.15)$$

$$\liminf \frac{1}{u_n} \log P_n(G) \geq -\inf\{I(x); x \in G\}, \quad (5.16)$$

for every closed set  $F \subset \mathcal{X}$  and every open set  $G \subset \mathcal{X}$ . The rate function  $I$  is called good if its level sets are compact. More generally, a sequence of  $\mathcal{X}$ -valued random variables is said to satisfy the LDP if their distributions satisfy the LDP. From now on, we work with  $\mathcal{X} = \mathcal{M}_1(\mathbb{T})$  whose compacity makes our task simpler. In fact, according to [10] Theorem 4.1.11, in this case the LDP is equivalent to the following property: for every  $\mu \in \mathcal{M}_1(\mathbb{T})$

$$-I(\mu) = \lim_{\varepsilon \downarrow 0} \limsup_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)) \quad (5.17)$$

$$= \lim_{\varepsilon \downarrow 0} \liminf_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)) \quad (5.18)$$

where  $B(\mu, \varepsilon)$  is the open ball of radius  $\varepsilon$ , centered at  $\mu$

$$B(\mu, \varepsilon) = \{\nu : d_L(\mu, \nu) < \varepsilon\}.$$

It is easy to see that the latter is equivalent to the pair of inequalities

$$\lim_{\varepsilon \downarrow 0} \limsup_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)) \leq -I(\mu) \quad (5.19)$$

$$\lim_{\varepsilon \downarrow 0} \limsup_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)) \geq -I(\mu). \quad (5.20)$$

For the sake of completeness, let us explain shortly why (5.19) and (5.20) lead to (5.15) and (5.16), respectively. From the one hand, every closed (hence compact) set in  $\mathcal{M}_1(\mathbb{T})$  may be recovered by a finite number of balls and applying (5.19) for all these balls leads to (5.15). On the other hand, every open set contains open balls and it remains to optimize with respect to the centers of the balls.

Our large deviations result follows the way initiated by the pioneer paper of Ben Arous and Guionnet ([5]) and continued by Hiai and Petz ([16],[18]).

We work with the set  $\mathcal{M}_1(\mathbb{T})$  of probability measures on the unit circle. For  $\mu \in \mathcal{M}_1(\mathbb{T})$ , the Voiculescu entropy is

$$\Sigma(\mu) = \iint \log |\zeta - \zeta'| d\mu(\zeta) d\mu(\zeta'). \quad (5.21)$$

With a different sign it is the logarithmic energy of  $\mu$  (see Chapter 5.3 in [17]). We also define the potential

$$Q_d(\zeta) = \begin{cases} -2(\Re d) \log(2 \sin \frac{\theta}{2}) - (\Im d)(\theta - \pi) & \text{if } \zeta = e^{i\theta}, \theta \in (0, 2\pi) \\ \infty & \text{if } \zeta = 1 \text{ and } \Re d > 0 \\ -|\Im d|\pi & \text{if } \zeta = 1 \text{ and } \Re d = 0 \end{cases} \quad (5.22)$$

It should be noticed that  $Q_d$  is a lower semi continuous function (l.s.c.). The main result of this section is an extension of the theorem of Hiai and Petz (Theorem 5.4.10 in [17]), which corresponds to the case  $d = 0 + a$  continuous potential.

**Theorem 5.5.** *Let  $d \in \mathbb{C}$  with  $\Re d \geq 0$ . For  $n \in \mathbb{N}$  and  $(\zeta_1, \dots, \zeta_n) \in \mathbb{T}^n$  let*

$$h(\zeta_1, \dots, \zeta_n) := \prod_{k=1}^n (1 - \zeta_k)^{\bar{d}\beta'n} (1 - \bar{\zeta}_k)^{d\beta'n} \prod_{j < k} |\zeta_j - \zeta_k|^{2\beta'}$$

and let  $\mathbb{P}_d^{(n)}$  be the distribution on  $\mathbb{T}^n$  having the density

$$\frac{1}{\mathcal{Z}_d(n)} h(\zeta_1, \dots, \zeta_n). \quad (5.23)$$

where  $\mathcal{Z}_d(n)$  is the normalization constant.

(1) We have

$$\lim_{n \rightarrow \infty} \frac{1}{n^2 \beta'} \log \mathcal{Z}_d(n) = B(d) \quad (5.24)$$

where

$$B(d) = \int_0^1 x \log \frac{x(x + \Re d)}{|x + d|^2} dx.$$

(2) The sequence of distributions of

$$\mu_{\text{esd}}^{(n)} = \frac{\delta_{\zeta_1} + \cdots + \delta_{\zeta_n}}{n}$$

under  $\mathbb{P}_d^{(n)}$  satisfies the LDP in  $\mathcal{M}_1(\mathbb{T})$  at scale  $\beta' n^2$  with (good) rate function

$$I_d(\mu) = -\Sigma(\mu) + \int_{\mathbb{T}} Q_d(\zeta) d\mu(\zeta) + B(d). \quad (5.25)$$

(3) The rate function vanishes only at  $\mu = \mu_d$ .

*Proof.* (1) An exact expression of  $\mathcal{Z}(n)$  is obtained using the following lemma, whose proof is postponed to the end of this subsection.

**Lemma 5.6.** *If we define*

$$\mathcal{Z}_{s,t}(n) = \int_{[0, 2\pi)^n} \prod_{k=1}^n (1 - e^{i\theta_k})^s (1 - e^{-i\theta_k})^t \prod_{j < k} |e^{i\theta_j} - e^{i\theta_k}|^\beta d\theta_1 \dots d\theta_n,$$

then we have

$$\mathcal{Z}_{s,t}(n) = \frac{\Gamma(\beta' n + 1)}{(\Gamma(\beta' + 1))^n} \prod_0^{n-1} \frac{\Gamma(\beta' j + 1) \Gamma(\beta' j + 1 + s + t)}{\Gamma(\beta' j + 1 + s) \Gamma(\beta' j + 1 + t)}. \quad (5.26)$$

We have  $\mathcal{Z}_d(n) = \mathcal{Z}_{\bar{d}\beta'n, d\beta'n}(n)$  and then,

$$\begin{aligned} \log \mathcal{Z}_d(n) &= \log \Gamma(\beta' n + 1) - n \log \Gamma(\beta' + 1) + \sum_{j=0}^{n-1} \log \Gamma(\beta' j + 1) \\ &+ \sum_{j=0}^{n-1} [\log \Gamma(\beta' j + 1 + 2\Re d \beta' n) - 2\Re d \log \Gamma(\beta' j + 1 + d\beta' n)]. \end{aligned}$$

From the Binet formula (Abramowitz and Stegun [1] or Erdélyi et al. [2] p.21), we have for  $\Re x > 0$

$$\log \Gamma(x) = (x - \frac{1}{2}) \log x - x + \frac{1}{2} \log(2\pi) + \int_0^\infty f(s) e^{-sx} ds. \quad (5.27)$$

where the function  $f$  is defined by

$$f(s) = \left[ \frac{1}{2} - \frac{1}{s} + \frac{1}{e^s - 1} \right] \frac{1}{s} = 2 \sum_{k=1}^{\infty} \frac{1}{s^2 + 4\pi^2 k^2},$$

and satisfies for every  $s \geq 0$

$$0 < f(s) \leq f(0) = 1/12, \quad 0 < \left( sf(s) + \frac{1}{2} \right) < 1.$$

Using (5.27), a straightforward study of Riemann sums gives

$$\frac{1}{\beta'n^2} \log \mathcal{Z}_d(n) \rightarrow \int_0^1 x \log \frac{x(x + \Re \mathbf{e} d)}{|x + d|^2} dx = B(d). \quad (5.28)$$

(2) The proof is based on the explicit form of the joint eigenvalue density. We follow the lines of [5], [16], [18] and [17]. We skip the index  $\mathbf{e}sd$  for notational convenience. Our goal is the proof of the two inequalities (cf. (5.19) and (5.20)), which hold for every  $\mu \in \mathcal{M}_1(\mathbb{T})$ :

$$\lim_{\varepsilon \rightarrow 0} \limsup_n \frac{1}{\beta'n^2} \log \mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \leq -I_d(\mu) \quad (5.29)$$

$$\lim_{\varepsilon \rightarrow 0} \liminf_n \frac{1}{\beta'n^2} \log \mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \geq -I_d(\mu) \quad (5.30)$$

We rest on the LDP known in the case of  $d = 0$ , and use the general method which consists in estimating the Radon-Nikodym derivative ruling the change of probability. The key formula is

$$\mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) = \frac{\mathcal{Z}_0(n)}{\mathcal{Z}_d(n)} \mathbb{E}_0^{(n)} \left[ \mathbf{1}_{\mu^{(n)} \in B(\mu, \varepsilon)} e^{-n^2 \beta' \int Q_d d\mu^{(n)}} \right]. \quad (5.31)$$

*The upper bound (5.29).* Let us first assume  $\Re \mathbf{e} d > 0$ , so that  $Q_d(\zeta) \rightarrow \infty$  as  $\zeta \rightarrow 1$ . For  $R > 0$  we consider the cutoff  $Q^R = \min(Q_d, R)$ . Since  $Q^R$  is continuous, the mapping  $\nu \in \mathcal{M}_1(\mathbb{T}) \mapsto \int Q^R d\nu$  is continuous, so

$$\inf_{B(\mu, \varepsilon)} \int Q^R d\nu \geq \int Q^R d\mu - r_1(\varepsilon, R) \quad (5.32)$$

with  $\lim_{\varepsilon} r_1(\varepsilon, R) = 0$ . Since  $Q_d \geq Q^R$  we get

$$\begin{aligned} \frac{1}{\beta'n^2} \log \mathbb{E}_0^{(n)} \left[ \mathbf{1}_{\mu^{(n)} \in B(\mu, \varepsilon)} e^{-n^2 \beta' \int Q_d d\mu^{(n)}} \right] &\leq \frac{1}{\beta'n^2} \log \mathbb{P}_0^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \\ &\quad - \int Q^R d\mu + r_1(\varepsilon, R) \end{aligned} \quad (5.33)$$

Thanks to the LDP known for  $d = 0$ , we can take limsup in  $n$  and then limit in  $\varepsilon$  and obtain

$$\lim_{\varepsilon} \limsup_n \frac{1}{\beta'n^2} \log \mathbb{E}_0^{(n)} \left[ \mathbf{1}_{\mu^{(n)} \in B(\mu, \varepsilon)} e^{-n^2 \beta' \int Q_d d\mu^{(n)}} \right] \leq \Sigma(\mu) - \int Q^R d\mu,$$

and thanks to (5.31) and (5.24)

$$\lim_{\varepsilon} \limsup_n \frac{1}{\beta'n^2} \log \mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \leq \Sigma(\mu) - \int Q^R d\mu - B(d).$$

By the monotone convergence theorem  $\lim_{R \rightarrow \infty} \int Q^R d\mu = \int Q_d d\mu$ , which proves (5.29).

If  $\Re d = 0$ , then  $Q_d$  is lower semi continuous and bounded, so that the mapping  $\nu \mapsto \int Q_d d\nu$  is lower semi continuous and (5.32) still holds with  $Q_d$  instead of  $Q^R$  and some  $r(\varepsilon)$  instead of  $r_1(\varepsilon, R)$ . The rest of the argument is the same as above.

*The lower bound (5.30).* If  $I_d(\mu) = \infty$ , the bound is trivial, so that we can assume that  $\mu$  has no atom. To overcome the problem of the singularity at 1, we use a classical approximation of  $\mu$  by a probability vanishing in a neighborhood of 1, i.e.

$$d\mu_M(\zeta) = \frac{\mathbf{1}_{|1-\zeta| \geq M^{-1}}}{\mu(|1-\zeta| \geq M^{-1})} d\mu(\zeta).$$

The benefit is that the mapping  $\nu \mapsto \int Q_d d\nu$  is continuous in a neighborhood of  $\mu_M$ . Choosing  $M$  large enough, we ensure that  $d_L(\mu_M, \mu) \leq \varepsilon/2$  and then, thanks to the triangle inequality

$$B(\mu, \varepsilon) \supset B(\mu_M, \varepsilon/2),$$

which leads to

$$\mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \geq \mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu_M, \varepsilon/2)).$$

We come back to (5.31) with  $\mu_M$  and  $\varepsilon/2$ . For  $\nu \in B(\mu_M, \varepsilon/2)$ , we have  $\int Q_d d\nu \leq \int Q_d d\mu_M + r_2(\varepsilon, M)$  where  $\lim_{\varepsilon} r_2(\varepsilon, M) = 0$ . We get

$$\begin{aligned} \frac{1}{\beta' n^2} \log \mathbb{E}_0^{(n)} \left[ \mathbf{1}_{\mu^{(n)} \in B(\mu, \varepsilon)} e^{-n^2 \beta' \int Q_d d\mu^{(n)}} \right] &\geq \frac{1}{\beta' n^2} \mathbb{P}_0^{(n)}(\mu^{(n)} \in B(\mu_M, \varepsilon/2)) \\ &\quad - \int Q_d d\mu_M - r_2(\varepsilon, M) \end{aligned}$$

Again, we take  $\liminf$  in  $n$  and  $\lim$  in  $\varepsilon$ , and use successively the LDP for  $d = 0$ , (5.31) and (5.24) to obtain:

$$\lim_{\varepsilon} \limsup_n \frac{1}{\beta' n^2} \log \mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \geq \Sigma(\mu_M) - \int Q_d d\mu_M - B(d).$$

Now, since  $I(\mu) < \infty$  and since  $Q_d$  and  $\Sigma$  are bounded below, the monotone convergence theorem yields

$$\lim_{M \rightarrow \infty} \Sigma(\mu_M) = \Sigma(\mu) \quad , \quad \lim_{M \rightarrow \infty} \int Q_d d\mu_M = \int Q_d d\mu.$$

This ends the proof of (5.30).

(3) The uniqueness of the minimizer is a direct consequence of the strict convexity of  $I$  which comes from the strict concavity of  $\Sigma$ .

We do not give a self contained proof of the identity of the minimizer, but rather use a probabilistic argument. On the one hand in Theorem 5.4 we proved that  $\mu_{\text{esd}}^{(n)}$  converges weakly in probability to  $\mu_d$ , and on the other hand the LDP combined with the uniqueness of the minimizer imply that  $\mu_{\text{esd}}^{(n)}$  converges weakly in probability to this minimizer. This completes the proof.  $\square$

*Proof of Lemma 5.6.* We have

$$\frac{\mathcal{Z}_{s,t}(n)}{\mathcal{Z}_{0,0}(n)} = \mathbb{E}(\det(\text{Id} - U)^s \det(\text{Id} - \bar{U})^t)$$

where the mean is taken under the  $\text{CJ}_{0,\beta}^{(n)}$  distribution. Under this distribution,  $\det(\text{Id} - U)$  has the same law as the product of independent variables  $1 - \bar{\alpha}_k$ , where  $\alpha_k$  is  $\nu_{\beta(n-k-1)+1}$  distributed. We get

$$\mathbb{E}(\det(\text{Id} - U)^s \det(\text{Id} - \bar{U})^t) = \prod_{j=0}^{n-1} \mathbb{E}(1 - \bar{\alpha}_j)^s (1 - \alpha_j)^t$$

From (3.3) we get

$$\frac{\mathcal{Z}_{s,t}(n)}{\mathcal{Z}_{0,0}(n)} = \prod_0^{n-1} \frac{\Gamma(\beta'j + 1)\Gamma(\beta'j + 1 + s + t)}{\Gamma(\beta'j + 1 + s)\Gamma(\beta'j + 1 + t)}$$

Besides, Lemma 4.4 in [24] gives

$$\mathcal{Z}_{0,0}(n) = \frac{\Gamma(\beta'n + 1)}{(\Gamma(\beta' + 1))^n}. \quad \square$$

#### REFERENCES

- [1] M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York (1972), 9th edition, 258-259.
- [2] A. Erdelyi, W. Magnus, F. Oberhettinger and F.G. Tricomi, Higher transcendental functions, Krieger, New-York (1981), I, 15-20.
- [3] G. Ammar, W. Gragg, and L. Reichel, Constructing a unitary Hessenberg matrix from spectral data, Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms (Leuven, 1988), pp. 385-395, NATO Adv. Sci. Inst. Ser. F Comput. Systems Sci., 70.
- [4] G.E. Andrews, R.A. Askey, R. Roy, Special functions, Encyclopedia of Mathematics and its Applications, 71, Cambridge University Press, Cambridge, 1999.
- [5] G. Ben Arous, and A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy, Probab. Theory Related Fields, (108) (1997), 517-542.
- [6] A. Borodin, G. Olshanski, Infinite Random Matrices and Ergodic Measures, Comm. Math. Phys. 203 (2001), 87-123.
- [7] P. Bourgade, C.P. Hughes, A. Nikeghbali, M. Yor, The characteristic polynomial of a random unitary matrix: a probabilistic approach, Duke Mathematical Journal vol 145, Number 1, 45-69, 2008.
- [8] P. Bourgade, A. Nikeghbali, A. Rouault, Hua-Pickrell measures on general compact groups, preprint, arXiv:0712.0848.
- [9] M. J. Cantero, L. Moral, and L. Velazquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, Linear Algebra Appl. 362 (2003), 29-56.
- [10] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications, Springer 1998 2nd edition.
- [11] I. Dumitriu and A. Edelman, Matrix models for beta ensembles. J. Math. Phys. 43 (2002), 5830-5847.
- [12] F. Dyson, Statistical theory of the energy levels of complex systems. I, II, and III. J. Math. Phys. 3 (1962), 140-156, 157-165, and 166-175.
- [13] P. J. Forrester, Log-gases and random matrices, book online.



- [14] P.J. Forrester and E.M. Rains, Jacobians and rank 1 perturbations relating to unitary Hessenberg matrices, *International Mathematics Research Notices*, **2006** (2006).
- [15] P.J. Forrester and N.S. Witte, Application of the  $\tau$ -function theory of Painleve equations to random matrices: Pvi, the JUE, CyUE, cJUE and scaled limits, *Nagoya Math. J.* 174 (2004), p.29-114.
- [16] F. Hiai and D. Petz, A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices, *Ann. Inst. H. Poincaré Probab. Statist.*, 36, (2000), 71-85.
- [17] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, Mathematical Surveys and Monographs, 77 (2000), Amer. Math. Soc., Providence.
- [18] F. Hiai and D. Petz, Large deviations for functions of two random projection matrices, *Acta Sci. Math. (Szeged)* 72 (2006) 581-609.
- [19] L. K. Hua, *Harmonic analysis of functions of several complex variables in the classical domains*, Chinese edition: Science Press, Peking, 1958; Russian edition: IL, Moscow, 1959; English edition: Transl. Math. Monographs 6, Amer. Math. Soc., 1963.
- [20] J.P. Keating and N.C. Snaith, Random Matrix Theory and  $\zeta(1/2 + it)$ , *Comm. Math. Phys.* 214, p 57-89, 2000.
- [21] J.P. Keating and N.C. Snaith, Random matrix theory and  $L$ -functions at  $s = 1/2$ , *Comm. Math. Phys.* **214** (2000) 91–110.
- [22] S.V. Khrushchev, Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in  $L^2(\mathbb{T})$ , *J. Approx. Theory* 108 (2) (2001), p. 161-248.
- [23] S.V. Khrushchev, Classification theorems for general orthogonal polynomials on the unit circle, *J. Approx. Theory* (116) (2) 2002 268–342.
- [24] R. Killip and I. Nenciu, Matrix models for circular ensembles, *International Mathematics Research Notices*, vol. 2004, no. 50, pp. 2665-2701, 2004.
- [25] R. Killip and I. Nenciu, CMV : The unitary analogue of Jacobi matrices, *Communications on Pure and Applied Mathematics*, Volume 60, Issue 8, 1148-1188, 2006.
- [26] R. Killip and M. Stoiciu, Eigenvalue statistics for CMV matrices: from Poisson to clock via  $C\beta E$ , preprint, arXiv:math-ph/0608002v1, 2006.
- [27] M. L. Mehta, *Random matrices*. Second edition. Academic Press, Inc., Boston, MA, 1991.
- [28] F. Mezzadri, N.C. Snaith (editors), *Recent Perspectives in Random Matrix Theory and Number Theory*, London Mathematical Society Lecture Note Series 322 (CUP), (2005).
- [29] Yu. A. Neretin, Hua type integrals over unitary groups and over projective limits of unitary groups, *Duke Math. J.* 114 (2002), 239-266.
- [30] D. Pickrell, Measures on infinite-dimensional Grassmann manifolds, *J. Func. Anal.* 70 (1987), no. 2, 323-356.
- [31] D. Pickrell, Mackey analysis of infinite classical motion groups, *Pacific J. Math.* 150 (1991), 139-166.
- [32] S. Port, *Theoretical Probability for Applications*, Wiley, 1994.
- [33] B. Simon, CMV matrices: Five years later, *J. Comput. Appl. Math.*, 208: 120-154, 2007.
- [34] B. Simon, OPUC one one foot, *Bulletin of the American Mathematical Society*, Vol. 42, Number 4, 431-460, 2005.
- [35] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Publications, Vol. 54.1, American Mathematical Society, Providence, RI, 2005.
- [36] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Publications, Vol. 54.2, American Mathematical Society, Providence, RI, 2005.
- [37] N.S. Witte and P.J. Forrester, Gap probabilities in the finite and scaled Cauchy random matrix ensembles, *Nonlinearity*, Vol. 13, 1965-1986, 2000.

ENST, 46 RUE BARRAULT, 75634 PARIS CEDEX 13. UNIVERSITÉ PARIS 6, LPMA,  
175, RUE DU CHEVALERET F-75013 PARIS.

*E-mail address:* `bourgade@enst.fr`

INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH, WINTERTHURERSTRASSE 190,  
CH-8057 ZÜRICH, SWITZERLAND

*E-mail address:* `ashkan.nikeghbali@math.uzh.ch`

UNIVERSITÉ VERSAILLES-SAINT QUENTIN, LMV, BÂTIMENT FERMAT, 45 AVENUE DES  
ETATS-UNIS, 78035 VERSAILLES CEDEX

*E-mail address:* `alain.rouault@math.uvsq.fr`