Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Approximation Theory 139 (2006) 29–64

JOURNAL OF Approximation Theory

www.elsevier.com/locate/jat

# The statistical distribution of the zeros of random paraorthogonal polynomials on the unit circle

Mihai Stoiciu

*Mathematics 253-37, California Institute of Technology, Pasadena, CA 91125, USA*

Received 10 December 2004; accepted 13 April 2005

Communicated by Paul Nevai
Available online 11 July 2005

**Abstract**

The orthogonal polynomials on the unit circle are defined by the recurrence relation

$$\Phi_{k+1}(z) = z\Phi_k(z) - \overline{\alpha}_k \Phi_k^*(z), \quad k \geqslant 0, \quad \Phi_0 = 1,$$

where $\alpha_k \in \mathbb{D}$ for any $k \geqslant 0$. If we consider $n$ complex numbers $\alpha_0, \alpha_1, \ldots, \alpha_{n-2} \in \mathbb{D}$ and $\alpha_{n-1} \in \partial\mathbb{D}$, we can use the previous recurrence relation to define the monic polynomials $\Phi_0, \Phi_1, \ldots, \Phi_n$. The polynomial $\Phi_n(z) = \Phi_n(z; \alpha_0, \ldots, \alpha_{n-2}, \alpha_{n-1})$ obtained in this way is called the paraorthogonal polynomial associated to the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$.

We take $\alpha_0, \alpha_1, \ldots, \alpha_{n-2}$ i.i.d. random variables distributed uniformly in a disk of radius $r < 1$ and $\alpha_{n-1}$ another random variable independent of the previous ones and distributed uniformly on the unit circle. For any $n$ we will consider the random paraorthogonal polynomial $\Phi_n(z) = \Phi_n(z; \alpha_0, \ldots, \alpha_{n-2}, \alpha_{n-1})$. The zeros of $\Phi_n$ are $n$ random points on the unit circle.

We prove that for any $e^{i\theta} \in \partial\mathbb{D}$ the distribution of the zeros of $\Phi_n$ in intervals of size $O(\frac{1}{n})$ near $e^{i\theta}$ is the same as the distribution of $n$ independent random points uniformly distributed on the unit circle (i.e., Poisson). This means that, for large $n$, there is no local correlation between the zeros of the considered random paraorthogonal polynomials.


*MSC:* primary 42C05; secondary 82B44

*Keywords:* Zeros of orthogonal polynomials; Random verblunsky coefficients; Random CMV matrices

*E-mail address:* mihai@caltech.edu.

0021-9045/$ - see front matter 
doi:10.1016/j.jat.2005.04.001

## 1. Introduction

In this paper, we study the statistical distribution of the zeros of paraorthogonal polynomials on the unit circle. In order to introduce and motivate these polynomials, we will first review a few aspects of the standard theory. Complete references for both the classical and the spectral theory of orthogonal polynomials on the unit circle are Simon [29,30].

One of the central results in this theory is Verblunsky's theorem, which states that there is a one–one and onto map $\mu \to \{\alpha_n\}_{n\geqslant 0}$ from the set of nontrivial (i.e., not supported on a finite set) probability measures on the unit circle and sequence of complex numbers $\{\alpha_n\}_{n\geqslant 0}$ with $|\alpha_n| < 1$ for any $n$. The correspondence is given by the recurrence relation obeyed by orthogonal polynomials on the unit circle. Thus, if we apply the Gram–Schmidt procedure to the sequence of polynomials $1, z, z^2, \ldots \in L^2(\partial\mathbb{D}, d\mu)$, the polynomials obtained $\Phi_0(z, d\mu), \Phi_1(z, d\mu), \Phi_2(z, d\mu) \ldots$ obey the recurrence relation

$$\Phi_{k+1}(z, d\mu) = z\Phi_k(z, d\mu) - \overline{\alpha}_k \Phi_k^*(z, d\mu), \quad k \geqslant 0, \tag{1.1}$$

where for $\Phi_k(z, d\mu) = \sum_{j=0}^{k} b_{k,j}\, z^j$, the reversed polynomial $\Phi_k^*$ is given by $\Phi_k^*(z, d\mu) = \sum_{j=0}^{k} \overline{b}_{k,k-j}\, z^j$. The numbers $\alpha_k$ from (1.1) obey $|\alpha_k| < 1$ and, for any $k$, the zeros of the polynomial $\Phi_{k+1}(z, d\mu)$ lie inside the unit disk.

If, for a fixed $n$, we take $\alpha_0, \alpha_1, \ldots, \alpha_{n-2} \in \mathbb{D}$ and $\alpha_{n-1} = \beta \in \partial\mathbb{D}$ and we use the recurrence relations (1.1) to define the polynomials $\Phi_0, \Phi_1, \ldots, \Phi_{n-1}$, then the zeros of the polynomial

$$\Phi_n(z, d\mu, \beta) = z\Phi_{n-1}(z, d\mu) - \overline{\beta}\, \Phi_{n-1}^*(z, d\mu) \tag{1.2}$$

are simple and situated on the unit circle. These polynomials (obtained by taking the last Verblunsky coefficient on the unit circle) are called paraorthogonal polynomials and were analyzed in [17,20]; see also Chapter 2 in Simon [29].

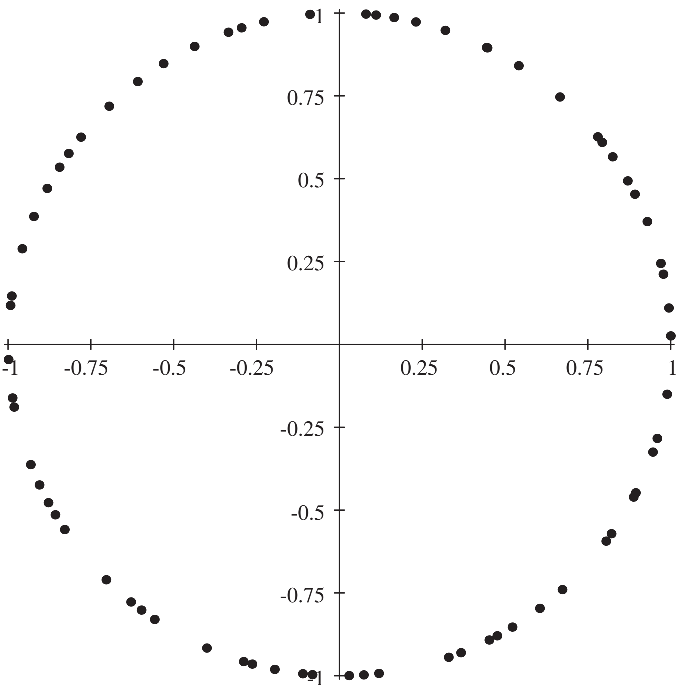
For any $n$, we will consider random Verblunsky coefficients by taking $\alpha_0, \alpha_1, \ldots, \alpha_{n-2}$ to be i.i.d. random variables distributed uniformly in a disk of fixed radius $r < 1$ and $\alpha_{n-1}$ another random variable independent of the previous ones and distributed uniformly on the unit circle. Following the procedure mentioned before, we will get a sequence of random paraorthogonal polynomials $\{\Phi_n = \Phi_n(z, d\mu, \beta)\}_{n\geqslant 0}$. For any $n$, the zeros of $\Phi_n$ are $n$ random points on the unit circle. Let us consider

$$\zeta^{(n)} = \sum_{k=1}^{n} \delta_{z_k^{(n)}}, \tag{1.3}$$

where $z_1^{(n)}, z_2^{(n)}, \ldots, z_n^{(n)}$ are the zeros of the polynomial $\Phi_n$. Let us also fix a point $e^{i\theta} \in \partial\mathbb{D}$. We will prove that the distribution of the zeros of $\Phi_n$ on intervals of length $O(\frac{1}{n})$ situated near $e^{i\theta}$ is the same as the distribution of $n$ independent random points uniformly distributed in the unit circle (i.e., Poisson).

A collection of random points on the unit circle is sometimes called a point process on the unit circle. Therefore, a reformulation of this problem can be: the limit of the sequence point process $\{\zeta^{(n)}\}_{n\geqslant 0}$ on a fine scale (of order $O(\frac{1}{n})$) near a point $e^{i\theta}$ is a Poisson point process.

This result is illustrated by the following generic plot of the zeros of random paraorthogonal polynomials:




This Mathematical plot represents the zeros of a paraorthogonal polynomial of degree 71 obtained by randomly taking $\alpha_0, \alpha_1, \ldots, \alpha_{69}$ from the uniform distribution on the disk centered at the origin of radius $\frac{1}{2}$ and random $\alpha_{70}$ from the uniform distribution on the unit circle. On a fine scale we can observe some clumps, which suggests the Poisson distribution.

Similar results appeared in the mathematical literature for the case of random Schrödinger operators; see [25,26]. The study of the spectrum of random Schrödinger operators and of the distribution of the eigenvalues was initiated by the very important paper of Anderson [4], who showed that certain random lattices exhibit absence of diffusion. Rigorous mathematical proofs of the Anderson localization were given by Goldsheid–Molchanov–Pastur [16] for one-dimensional models and by Fröhlich–Spencer [13] for multidimensional Schrödinger operators. Several other proofs, containing improvements and simplifications, were published later. We will only mention here Aizenman–Molchanov [2] and Simon–Wolff [33], which are relevant for our approach. In the case of the unit circle, similar localization results were obtained by Teplyaev [35] and by Golinskii–Nevai [18].

In addition to the phenomenon of localization, one can also analyze the local structure of the spectrum. It turns out that there is no repulsion between the energy levels of the Schrödinger operator. This was shown by Molchanov [26] for a model of the one-dimensional Schrödinger operator studied by the Russian school. The case of the multidimensional discrete Schrödinger operator was analyzed by Minami [25]. In both cases the authors proved that the statistical distribution of the eigenvalues converges locally to a stationary Poisson point process. This means that there is no correlation between eigenvalues.

We will prove a similar result on the unit circle. For any probability measure $d\mu$ on the unit circle, we denote by $\{\chi_0, \chi_1, \chi_2, \ldots\}$ the basis of $L^2(\partial\mathbb{D}, d\mu)$ obtained from $\{1, z, z^{-1}, z^2, z^{-2}, \ldots\}$ by applying the Gram–Schmidt procedure. The matrix representation of the operator $f(z) \to zf(z)$ on $L^2(\partial\mathbb{D}, d\mu)$ with respect to the basis $\{\chi_0, \chi_1, \chi_2, \ldots\}$ is a five-diagonal matrix of the form:

$$\mathcal{C} = \begin{pmatrix} \bar{\alpha}_0 & \bar{\alpha}_1\rho_0 & \rho_1\rho_0 & 0 & 0 & \ldots \\ \rho_0 & -\bar{\alpha}_1\alpha_0 & -\rho_1\alpha_0 & 0 & 0 & \ldots \\ 0 & \bar{\alpha}_2\rho_1 & -\bar{\alpha}_2\alpha_1 & \bar{\alpha}_3\rho_2 & \rho_3\rho_2 & \ldots \\ 0 & \rho_2\rho_1 & -\rho_2\alpha_1 & -\bar{\alpha}_3\alpha_2 & -\rho_3\alpha_2 & \ldots \\ 0 & 0 & 0 & \bar{\alpha}_4\rho_3 & -\bar{\alpha}_4\alpha_3 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \tag{1.4}$$

($\alpha_0, \alpha_1, \ldots$ are the Verblunsky coefficients associated to the measure $\mu$, and for any $n \geqslant 0$, $\rho_n = \sqrt{1 - |\alpha_n|^2}$). This matrix representation is a recent discovery of Cantero et al. [6]. The matrix $\mathcal{C}$ is called the CMV matrix and will be used in the study of the distribution of the zeros of the paraorthogonal polynomials.

Notice that if one of the $\alpha$'s is of absolute value 1, then the Gram–Schmidt process ends and the CMV matrix decouples. In our case, $|\alpha_{n-1}| = 1$, so $\rho_{n-1} = 0$ and therefore the CMV matrix decouples between $(n-1)$ and $n$ and the upper left corner is an $(n \times n)$ unitary matrix $\mathcal{C}^{(n)}$. The advantage of considering this matrix is that the zeros of $\Phi_n$ are exactly the eigenvalues of the matrix $\mathcal{C}^{(n)}$ (see, e.g., [29]). We will use some techniques from the spectral theory of the discrete Schrödinger operators to study the distribution of these eigenvalues, especially ideas and methods developed in [2,3,8,25,26,28]. However, our model on the unit circle has many different features compared to the discrete Schrödinger operator (perhaps the most important one is that we have to consider unitary operators on the unit circle instead of self-adjoint operators on the real line). Therefore, we will have to use new ideas and techniques that work for this situation.

The final goal is the following:

**Theorem 1.1.** *Consider the random polynomials on the unit circle given by the following recurrence relations*:

$$\Phi_{k+1}(z) = z\Phi_k(z) - \overline{\alpha}_k\Phi_k^*(z), \quad k \geqslant 0, \quad \Phi_0 = 1, \tag{1.5}$$

*where $\alpha_0, \alpha_1, \ldots, \alpha_{n-2}$ are i.i.d. random variables distributed uniformly in a disk of radius $r < 1$ and $\alpha_{n-1}$ is another random variable independent of the previous ones and uniformly distributed on the unit circle.*

*Consider the space $\Omega = \{\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, \alpha_{n-1}) \in D(0, r) \times D(0, r) \times \cdots \times D(0, r) \times \partial\mathbb{D}\}$ with the probability measure $\mathbb{P}$ obtained by taking the product of the uniform (Lebesgue) measures on each $D(0, r)$ and on $\partial\mathbb{D}$. Fix a point $e^{i\theta_0} \in \partial\mathbb{D}$ and let $\zeta^{(n)}$ be the point process defined by* (1.3).

*Then, on a fine scale (of order $\frac{1}{n}$) near $e^{i\theta_0}$, the point process $\zeta^{(n)}$ converges to the Poisson point process with intensity measure $n\,\frac{d\theta}{2\pi}$ (where $\frac{d\theta}{2\pi}$ is the normalized Lebesgue measure).*

*This means that for any fixed* $a_1 < b_1 \leqslant a_2 < b_2 \leqslant \cdots \leqslant a_m < b_m$ *and any nonnegative integers* $k_1, k_2, \ldots, k_m$, *we have*

$$\mathbb{P}\left(\zeta^{(n)}\left(e^{i(\theta_0+\frac{2\pi a_1}{n})}, e^{i(\theta_0+\frac{2\pi b_1}{n})}\right)\right.$$
$$\left. = k_1, \ldots, \zeta^{(n)}\left(e^{i(\theta_0+\frac{2\pi a_m}{n})}, e^{i(\theta_0+\frac{2\pi b_m}{n})}\right) = k_m\right)$$
$$\longrightarrow e^{-(b_1-a_1)}\frac{(b_1-a_1)^{k_1}}{k_1!}\ldots e^{-(b_m-a_m)}\frac{(b_m-a_m)^{k_m}}{k_m!} \tag{1.6}$$

*as* $n \to \infty$.

## 2. Outline of the proof

From now on we will work under the hypotheses of Theorem 1.1. We will study the statistical distribution of the eigenvalues of the random CMV matrices

$$\mathcal{C}^{(n)} = \mathcal{C}_\alpha^{(n)} \tag{2.1}$$

for $\alpha \in \Omega$ (with the space $\Omega$ defined in Theorem 1.1).

A first step in the study of the spectrum of random CMV matrix is proving the exponential decay of the fractional moments of the resolvent of the CMV matrix. These ideas were developed in the case of Anderson models by Aizenman–Molchanov [2] and by Aizenman et al. [3]. In the case of Anderson models, they provide a powerful method for proving spectral localization, dynamical localization, and the absence of level repulsion.

Before we state the Aizenman–Molchanov bounds, we have to make a few remarks on the boundary behavior of the matrix elements of the resolvent of the CMV matrix. For any $z \in \mathbb{D}$ and any $0 \leqslant k, l \leqslant (n-1)$, we will use the following notation:

$$F_{kl}(z, \mathcal{C}_\alpha^{(n)}) = \left[\frac{\mathcal{C}_\alpha^{(n)} + z}{\mathcal{C}_\alpha^{(n)} - z}\right]_{kl}. \tag{2.2}$$

As we will see in the next section, using properties of Carathéodory functions, we will get that for any $\alpha \in \Omega$, the radial limit

$$F_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)}) = \lim_{r\uparrow 1} F_{kl}(re^{i\theta}, \mathcal{C}_\alpha^{(n)}) \tag{2.3}$$

exists for Lebesgue almost every $e^{i\theta} \in \partial\mathbb{D}$ and $F_{kl}(\,\cdot\,, \mathcal{C}_\alpha^{(n)}) \in L^s(\partial\mathbb{D})$ for any $s \in (0,1)$. Since the distributions of $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ are rotationally invariant, we obtain that for any fixed $e^{i\theta} \in \partial\mathbb{D}$, the radial limit $F_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)})$ exists for almost every $\alpha \in \Omega$. We can also define

$$G_{kl}(z, \mathcal{C}_\alpha^{(n)}) = \left[\frac{1}{\mathcal{C}_\alpha^{(n)} - z}\right]_{kl} \tag{2.4}$$

and

$$G_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)}) = \lim_{r\uparrow 1} G_{kl}(re^{i\theta}, \mathcal{C}_\alpha^{(n)}). \tag{2.5}$$

Using the previous notation we have

**Theorem 2.1** (*Aizenman–Molchanov bounds for the resolvent of the CMV matrix*). *For the model considered in Theorem* 1.1 *and for any* $s \in (0,1)$, *there exist constants* $C_1, D_1 > 0$ *such that for any* $n > 0$, *any* $k, l$, $0 \leqslant k, l \leqslant n-1$ *and any* $e^{i\theta} \in \partial\mathbb{D}$, *we have*

$$\mathbb{E}\left(\left|F_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|^s\right) \leqslant C_1\, e^{-D_1|k-l|}, \tag{2.6}$$

*where* $\mathcal{C}_\alpha^{(n)}$ *is the* $(n \times n)$ *CMV matrix obtained for* $\alpha_0, \alpha_1, \ldots \alpha_{n-2}$ *uniformly distributed in* $D(0,r)$ *and* $\alpha_{n-1}$ *uniformly distributed in* $\partial\mathbb{D}$.

Using Theorem 2.1, we will then be able to control the structure of the eigenfunctions of the matrix $\mathcal{C}^{(n)}$.

**Theorem 2.2** (*The localized structure of the eigenfunctions*). *For the model considered in Theorem* 1.1, *the eigenfunctions of the random matrices* $\mathcal{C}^{(n)} = \mathcal{C}_\alpha^{(n)}$ *are exponentially localized with probability* 1, *that is exponentially small outside sets of size proportional to* $(\ln n)$. *This means that there exists a constant* $D_2 > 0$ *and for almost every* $\alpha \in \Omega$, *there exists a constant* $C_\alpha > 0$ *such that for any unitary eigenfunction* $\varphi_\alpha^{(n)}$, *there exists a point* $m(\varphi_\alpha^{(n)})$ $(1 \leqslant m(\varphi_\alpha^{(n)}) \leqslant n)$ *with the property that for any* $m$, $|m - m(\varphi_\alpha^{(n)})| \geqslant D_2 \ln(n+1)$, *we have*

$$|\varphi_\alpha^{(n)}(m)| \leqslant C_\alpha\, e^{-(4/D_2)\,|m - m(\varphi_\alpha^{(n)})|}. \tag{2.7}$$

The point $m(\varphi_\alpha^{(n)})$ will be taken to be the smallest integer where the eigenfunction $\varphi_\alpha^{(n)}(m)$ attains its maximum.

In order to obtain a Poisson distribution in the limit as $n \to \infty$, we will use the approach of Molchanov [26] and Minami [25]. The first step is to decouple the point process $\zeta^{(n)}$ into the direct sum of smaller point processes. We will do the decoupling process in the following way: for any positive integer $n$, let $\tilde{\mathcal{C}}^{(n)}$ be the CMV matrix obtained for the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_n$ with the additional restrictions $\alpha_{\left[\frac{n}{\ln n}\right]} = e^{i\eta_1}$, $\alpha_{2\left[\frac{n}{\ln n}\right]} = e^{i\eta_2}, \ldots, \alpha_n = e^{i\eta_{[\ln n]}}$, where $e^{i\eta_1}, e^{i\eta_2}, \ldots, e^{i\eta_{[\ln n]}}$ are independent random points uniformly distributed on the unit circle. Note that the matrix $\tilde{\mathcal{C}}^{(n)}$ decouples into the direct sum of $\approx [\ln n]$ unitary matrices $\tilde{\mathcal{C}}_1^{(n)}, \tilde{\mathcal{C}}_2^{(n)}, \ldots, \tilde{\mathcal{C}}_{[\ln n]}^{(n)}$. We should note here that the actual number of blocks $\tilde{\mathcal{C}}_i^{(n)}$ is slightly larger than $[\ln n]$ and that the dimension of one of the blocks (the last one) could be smaller than $\left[\frac{n}{\ln n}\right]$.

However, since we are only interested in the asymptotic behavior of the distribution of the eigenvalues, we can, without loss of generality, work with matrices of size $N = [\ln n]\left[\frac{n}{\ln n}\right]$. The matrix $\tilde{\mathcal{C}}^{(N)}$ is the direct sum of exactly $[\ln n]$ smaller blocks $\tilde{\mathcal{C}}_1^{(N)}, \tilde{\mathcal{C}}_2^{(N)}, \ldots, \tilde{\mathcal{C}}_{[\ln n]}^{(N)}$. We denote by $\zeta^{(N,p)} = \sum_{k=1}^{[n/\ln n]} \delta_{z_k^{(p)}}$ where $z_1^{(p)}, z_2^{(p)}, \ldots, z_{[n/\ln n]}^{(p)}$ are the eigenvalues of the matrix $\tilde{\mathcal{C}}_p^{(N)}$. The decoupling result is formulated in the following theorem:

**Theorem 2.3** (*Decoupling the point process*). *The point process $\zeta^{(N)}$ can be asymptotically approximated by the direct sum of point processes $\sum_{p=1}^{[\ln n]} \zeta^{(N,p)}$. In other words, the distribution of the eigenvalues of the matrix $\mathcal{C}^{(N)}$ can be asymptotically approximated by the distribution of the eigenvalues of the direct sum of the matrices $\tilde{\mathcal{C}}_1^{(N)}, \tilde{\mathcal{C}}_2^{(N)}, \ldots, \tilde{\mathcal{C}}_{[\ln n]}^{(N)}$.*

The decoupling property is the first step in proving that the statistical distribution of the eigenvalues of $\mathcal{C}^{(N)}$ is Poisson. In the theory of point processes (see, e.g., [7]), a point process obeying this decoupling property is called an infinitely divisible point process. In order to show that this distribution is Poisson on a scale of order $O(\frac{1}{n})$ near a point $e^{i\theta}$, we need to check two conditions:

$$\text{(i)} \quad \sum_{p=1}^{[\ln n]} \mathbb{P}\left(\zeta^{(N,p)}\left(A\left(N,\theta\right)\right) \geqslant 1\right) \to |A| \quad \text{as} \quad n \to \infty, \tag{2.8}$$

$$\text{(ii)} \quad \sum_{p=1}^{[\ln n]} \mathbb{P}\left(\zeta^{(N,p)}\left(A\left(N,\theta\right)\right) \geqslant 2\right) \to 0 \quad \text{as} \quad n \to \infty, \tag{2.9}$$

where for an interval $A = [a,b]$ we denote by $A(N,\theta) = (e^{i(\theta+\frac{2\pi a}{N})}, e^{i(\theta+\frac{2\pi b}{N})})$ and $|\cdot|$ is the Lebesgue measure (and we extend this definition to unions of intervals). The second condition shows that it is asymptotically impossible that any of the matrices $\tilde{\mathcal{C}}_1^{(N)}$, $\tilde{\mathcal{C}}_2^{(N)}, \ldots, \tilde{\mathcal{C}}_{[\ln n]}^{(N)}$ has two or more eigenvalues situated an interval of size $\frac{1}{N}$. Therefore, each of the matrices $\tilde{\mathcal{C}}_1^{(N)}, \tilde{\mathcal{C}}_2^{(N)}, \ldots, \tilde{\mathcal{C}}_{[\ln n]}^{(N)}$ contributes with at most one eigenvalue in an interval of size $\frac{1}{N}$. But the matrices $\tilde{\mathcal{C}}_1^{(N)}, \tilde{\mathcal{C}}_2^{(N)}, \ldots, \tilde{\mathcal{C}}_{[\ln n]}^{(N)}$ are decoupled, hence independent, and therefore we get a Poisson distribution. The condition (i) now gives Theorem 1.1.

The next four sections will contain the detailed proofs of these theorems.

## 3. Aizenman–Molchanov bounds for the resolvent of the CMV matrix

We will study the random CMV matrices defined in (2.1). We will analyze the matrix elements of the resolvent $(\mathcal{C}^{(n)} - z)^{-1}$ of the CMV matrix, or, what is equivalent, the matrix elements of

$$F(z, \mathcal{C}^{(n)}) = (\mathcal{C}^{(n)} + z)(\mathcal{C}^{(n)} - z)^{-1} = I + 2z\,(\mathcal{C}^{(n)} - z)^{-1} \tag{3.1}$$

(we consider $z \in \mathbb{D}$). More precisely, we will be interested in the expectations of the fractional moments of matrix elements of the resolvent. This method (sometimes called the fractional moments method) is useful in the study of the eigenvalues and of the eigenfunctions and was introduced by Aizenman and Molchanov in [2].

We will prove that the expected value of the fractional moment of the matrix elements of the resolvent decays exponentially (see (2.6)). The proof of this result is rather involved; the main steps will be

*Step* 1: The fractional moments $\mathbb{E}\left(\left|F_{kl}(z, \mathcal{C}_\alpha^{(n)})\right|^s\right)$ are uniformly bounded (Lemma 3.1).

*Step* 2: The fractional moments $\mathbb{E}\left(\left|F_{kl}(z,\mathcal{C}_\alpha^{(n)})\right|^s\right)$ converge to 0 uniformly along the rows (Lemma 3.6).

*Step* 3: The fractional moments $\mathbb{E}\left(\left|F_{kl}(z,\mathcal{C}_\alpha^{(n)})\right|^s\right)$ decay exponentially (Theorem 2.1).

We will now begin the analysis of $\mathbb{E}\left(\left|F_{kl}(z,\mathcal{C}_\alpha^{(n)})\right|^s\right)$.

It is not hard to see that $\mathrm{Re}\left[(\mathcal{C}^{(n)}+z)(\mathcal{C}^{(n)}-z)^{-1}\right]$ is a positive operator. This will help us prove

**Lemma 3.1.** *For any $s\in(0,1)$, any $k,l$, $1\leqslant k,l\leqslant n$, and any $z\in\mathbb{D}\cup\partial\mathbb{D}$, we have*

$$\mathbb{E}\left(\left|F_{kl}(z,\mathcal{C}_\alpha^{(n)})\right|^s\right)\leqslant C, \tag{3.2}$$

*where $C=\frac{2^{2-s}}{\cos\frac{\pi s}{2}}$.*

**Proof.** Let $F_\varphi(z)=(\varphi,(\mathcal{C}_\alpha^{(n)}+z)(\mathcal{C}_\alpha^{(n)}-z)^{-1}\varphi)$. Since $\mathrm{Re}\,F_\varphi\geqslant 0$, the function $F_\varphi$ is a Carathéodory function for any unit vector $\varphi$. Fix $\rho\in(0,1)$. Then, by a version of Kolmogorov's theorem (see [9] or [23]),

$$\int_0^{2\pi}\left|(\varphi,(\mathcal{C}_\alpha^{(n)}+\rho e^{i\theta})(\mathcal{C}_\alpha^{(n)}-\rho e^{i\theta})^{-1}\varphi)\right|^s\frac{d\theta}{2\pi}\leqslant C_1, \tag{3.3}$$

where $C_1=\frac{1}{\cos\frac{\pi s}{2}}$.

The polarization identity gives (assuming that our scalar product is antilinear in the first variable and linear in the second variable)

$$F_{kl}(\rho e^{i\theta},\mathcal{C}_\alpha^{(n)})=\frac{1}{4}\sum_{m=0}^{3}(-i)^m\left((\delta_k+i^m\delta_l),\ F(\rho e^{i\theta},\mathcal{C}_\alpha^{(n)})(\delta_k+i^m\delta_l)\right) \tag{3.4}$$

which, using the fact that $|a+b|^s\leqslant|a|^s+|b|^s$, implies

$$\left|F_{kl}(\rho e^{i\theta},\mathcal{C}_\alpha^{(n)})\right|^s\leqslant\frac{1}{2^s}\sum_{m=0}^{3}\left|\left(\frac{(\delta_k+i^m\delta_l)}{\sqrt{2}},\ F(\rho e^{i\theta},\mathcal{C}_\alpha^{(n)})\frac{(\delta_k+i^m\delta_l)}{\sqrt{2}}\right)\right|^s. \tag{3.5}$$

Using (3.3) and (3.5), we get, for any $\mathcal{C}_\alpha^{(n)}$,

$$\int_0^{2\pi}\left|F_{kl}(\rho e^{i\theta},\mathcal{C}_\alpha^{(n)})\right|^s\frac{d\theta}{2\pi}\leqslant C, \tag{3.6}$$

where $C=\frac{2^{2-s}}{\cos\frac{\pi s}{2}}$.

Therefore, after taking expectations and using Fubini's theorem,

$$\int_0^{2\pi} \mathbb{E}\left(\left|F_{kl}(\rho e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|^s\right) \frac{d\theta}{2\pi} \leqslant C. \tag{3.7}$$

The coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ define a measure $d\mu$ on $\partial\mathbb{D}$. Let us consider another measure $d\mu_\theta(e^{i\tau}) = d\mu(e^{i(\tau-\theta)})$. This measure defines Verblunsky coefficients $\alpha_{0,\theta}, \alpha_{1,\theta}, \ldots, \alpha_{n-1,\theta}$, a CMV matrix $\mathcal{C}_{\alpha,\theta}^{(n)}$, and orthonormal polynomials $\varphi_{0,\theta}, \varphi_{1,\theta}, \ldots, \varphi_{n-1,\theta}$. Using the results presented in Simon [29], for any $k$, $0 \leqslant k \leqslant n-1$,

$$\alpha_{k,\theta} = e^{-i(k+1)\theta}\alpha_k, \tag{3.8}$$

$$\varphi_{k,\theta}(z) = e^{ik\theta}\varphi_k(e^{-i\theta}z). \tag{3.9}$$

The relation (3.9) shows that for any $k$ and $\theta$, $\chi_{k,\theta}(z) = \lambda_{k,\theta}\, \chi_k(e^{-i\theta}z)$ where $|\lambda_{k,\theta}| = 1$.

Since $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ are independent and the distribution of each one of them is rotationally invariant, we have

$$\mathbb{E}\left(\left|F_{kl}(\rho e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|^s\right) = \mathbb{E}\left(\left|F_{kl}(\rho e^{i\theta}, \mathcal{C}_{\alpha,\theta}^{(n)})\right|^s\right). \tag{3.10}$$

But, using (3.8) and (3.9),

$$\begin{aligned}
F_{kl}(\rho e^{i\theta}, \mathcal{C}_{\alpha,\theta}^{(n)}) &= \int_{\partial\mathbb{D}} \frac{e^{i\tau} + \rho e^{i\theta}}{e^{i\tau} - \rho e^{i\theta}}\, \chi_{l,\theta}(e^{i\tau})\, \overline{\chi_{k,\theta}(e^{i\tau})}\, d\mu_\theta(e^{i\tau}) \\
&= \int_{\partial\mathbb{D}} \frac{e^{i\tau} + \rho e^{i\theta}}{e^{i\tau} - \rho e^{i\theta}}\, \chi_{l,\theta}(e^{i\tau})\, \overline{\chi_{k,\theta}(e^{i\tau})}\, d\mu(e^{i(\tau-\theta)}) \\
&= \int_{\partial\mathbb{D}} \frac{e^{i(\tau+\theta)} + \rho e^{i\theta}}{e^{i(\tau+\theta)} - \rho e^{i\theta}}\, \chi_{l,\theta}(e^{i(\tau+\theta)})\, \overline{\chi_{k,\theta}(e^{i(\tau+\theta)})}\, d\mu(e^{i\tau}) \\
&= \lambda_{l,\theta}\overline{\lambda}_{k,\theta} \int_{\partial\mathbb{D}} \frac{e^{i\tau} + \rho}{e^{i\tau} - \rho}\, \chi_l(e^{i\tau})\, \overline{\chi_k(e^{i\tau})}\, d\mu(e^{i\tau}) \\
&= \lambda_{l,\theta}\, \overline{\lambda}_{k,\theta}\, F_{kl}(\rho, \mathcal{C}_\alpha^{(n)}),
\end{aligned}$$

where $|\lambda_{l,\theta}\overline{\lambda}_{k,\theta}| = 1$.

Therefore the function $\theta \to \mathbb{E}\left(\left|F_{kl}(\rho e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|^s\right)$ is constant, so, using (3.7), we get

$$\mathbb{E}\left(\left|F_{kl}(\rho e^{i\theta}, \mathcal{C}_{\alpha,\theta}^{(n)})\right|^s\right) \leqslant C. \tag{3.11}$$

Since $\rho$ and $\theta$ are arbitrary, we now get the desired conclusion for any $z \in \mathbb{D}$.

Observe that by (3.4), $F_{kl}$ is a linear combination of Carathéodory functions. By Duren [9], any Carathéodory function is in $H^s(\mathbb{D})$ $(0 < s < 1)$ and therefore it has boundary values almost everywhere on $\partial\mathbb{D}$. Thus we get that, for any fixed $\alpha \in \Omega$ and for Lebesgue almost any $z = e^{i\theta} \in \partial\mathbb{D}$, the radial limit $F_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)})$ exists, where

$$F_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)}) = \lim_{\rho \uparrow 1} F_{kl}(\rho e^{i\theta}, \mathcal{C}_\alpha^{(n)}). \tag{3.12}$$

Also, by the properties of Hardy spaces, $F_{kl}(\,\cdot\,, \mathcal{C}_\alpha^{(n)}) \in L^s(\partial\mathbb{D})$ for any $s \in (0, 1)$. Since the distributions of $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ are rotationally invariant, we obtain that for any fixed $e^{i\theta} \in \partial\mathbb{D}$, the radial limit $F_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)})$ exists for almost every $\alpha \in \Omega$.

The relation (3.11) gives

$$\sup_{\rho\in(0,1)} \mathbb{E}\left(\left|F_{kl}(\rho e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|^s\right) \leqslant C. \tag{3.13}$$

By taking $\rho \uparrow 1$ and using Fatou's lemma we get

$$\mathbb{E}\left(\left|F_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|^s\right) \leqslant C. \qquad \square \tag{3.14}$$

Note that the argument from Lemma 3.1 works in the same way when we replace the unitary matrix $\mathcal{C}_\alpha^{(n)}$ with the unitary operator $\mathcal{C}_\alpha$ (corresponding to random Verblunsky coefficients uniformly distributed in $D(0, r)$), so we also have

$$\mathbb{E}\left(\left|F_{kl}(e^{i\theta}, \mathcal{C}_\alpha)\right|^s\right) \leqslant C \tag{3.15}$$

for any nonnegative integers $k, l$ and for any $e^{i\theta} \in \partial\mathbb{D}$.

The next step is to prove that the expectations of the fractional moments of the resolvent of $\mathcal{C}^{(n)}$ tend to zero on the rows. We will start with the following lemma suggested to us by Aizenman [1]:

**Lemma 3.2.** *Let $\{X_n = X_n(\omega)\}_{n\geqslant 0}$, $\omega \in \Omega$ be a family of positive random variables such that there exists a constant $C > 0$ such that $\mathbb{E}(X_n) < C$ and, for almost any $\omega \in \Omega$, $\lim_{n\to\infty} X_n(\omega) = 0$. Then, for any $s \in (0, 1)$,*

$$\lim_{n\to\infty} \mathbb{E}(X_n^s) = 0. \tag{3.16}$$

**Proof.** Let $\varepsilon > 0$ and let $M > 0$ such that $M^{s-1} < \varepsilon$. Observe that if $X_n(\omega) > M$, then $X_n^s(\omega) < M^{s-1} X_n(\omega)$. Therefore

$$X_n^s(\omega) \leqslant X_n^s(\omega)\, \chi_{\{\omega;\, X_n(\omega)\leqslant M\}}(\omega) + M^{s-1} X_n(\omega). \tag{3.17}$$

Clearly, $\mathbb{E}(M^{s-1}X_n) \leqslant \varepsilon C$ and, using dominated convergence,

$$\mathbb{E}(X_n^s\, \chi_{\{\omega;\, X_n(\omega)\leqslant M\}}) \to 0 \quad \text{as} \quad n \to \infty. \tag{3.18}$$

We immediately get that for any $\varepsilon > 0$ we have

$$\limsup_{n\to\infty} \mathbb{E}(X_n^s) \leqslant \mathbb{E}(X_n^s\, \chi_{\{\omega;\, X_n(\omega)\leqslant M\}}) + \varepsilon C \tag{3.19}$$

so we can conclude that (3.16) holds. $\square$

We will use Lemma 3.2 to prove that for any fixed $j$, $\mathbb{E}\left(\left|F_{j,j+k}(e^{i\theta}, \mathcal{C}_\alpha)\right|^s\right)$ and $\mathbb{E}\left(\left|F_{j,j+k}(e^{i\theta}, \mathcal{C}_\alpha^n)\right|^s\right)$ converge to 0 as $k \to \infty$. From now on, it will be more convenient to work with the resolvent $G$ instead of the Carathéodory function $F$.

**Lemma 3.3.** *Let $\mathcal{C} = \mathcal{C}_\alpha$ be the random CMV matrix associated to a family of Verblunsky coefficients $\{\alpha_n\}_{n\geqslant 0}$ with $\alpha_n$ i.i.d. random variables uniformly distributed in a disk $D(0, r)$, $0 < r < 1$. Let $s \in (0, 1)$, $z \in \mathbb{D} \cup \partial\mathbb{D}$, and $j$ a positive integer. Then we have*

$$\lim_{k\to\infty} \mathbb{E}\left(\left|G_{j,j+k}(z, \mathcal{C})\right|^s\right) = 0. \tag{3.20}$$

**Proof.** For any fixed $z \in \mathbb{D}$, the rows and columns of $G(z, \mathcal{C})$ are $l^2$ at infinity, hence converge to 0. Let $s' \in (s, 1)$. Then we get (3.20) applying Lemma 3.2 to the random variables $X_k = \left|G_{j,j+k}(z, \mathcal{C})\right|^{s'}$ and using the power $\frac{s}{s'} < 1$.

We will now prove (3.20) for $z = e^{i\theta} \in \partial\mathbb{D}$. In order to do this, we will have to apply the heavy machinery of transfer matrices and Lyapunov exponents developed in [30]. Thus, the transfer matrices corresponding to the CMV matrix are

$$T_n(z) = A(\alpha_n, z) \ldots A(\alpha_0, z), \tag{3.21}$$

where $A(\alpha, z) = (1 - |\alpha|^2)^{-1/2} \begin{pmatrix} z & -\overline{\alpha} \\ -\alpha z & 1 \end{pmatrix}$ and the Lyapunov exponent is

$$\gamma(z) = \lim_{n\to\infty} \frac{1}{n} \log \|T_n(z, \{\alpha_n\})\| \tag{3.22}$$

(provided this limit exists).

Observe that the common distribution $d\mu_\alpha$ of the Verblunsky coefficients $\alpha_n$ is rotationally invariant and

$$\int_{D(0,1)} - \log(1 - \omega)\, d\mu_\alpha(\omega) < \infty, \tag{3.23}$$

and

$$\int_{D(0,1)} - \log|\omega|\, d\mu_\alpha(\omega) < \infty. \tag{3.24}$$

Let us denote by $d\nu_N$ the density of eigenvalues measure and let $U^{d\nu_N}$ be the logarithmic potential of the measure $d\nu_N$, defined by

$$U^{d\nu_N}(e^{i\theta}) = \int_{\partial\mathbb{D}} \log \frac{1}{|e^{i\theta} - e^{i\tau}|}\, d\nu_N(e^{i\tau}). \tag{3.25}$$

By rotation invariance, we have $d\nu_N = \frac{d\theta}{2\pi}$ and therefore $U^{d\nu_N}$ is identically zero. Using results from [30], the Lyapunov exponent exists for every $z = e^{i\theta} \in \partial\mathbb{D}$ and the Thouless formula gives

$$\gamma(z) = -\tfrac{1}{2} \int_{D(0,1)} \log(1 - |\omega|^2)\, d\mu_\alpha(\omega). \tag{3.26}$$

By an immediate computation we get $\gamma(z) = \frac{r^2+(1-r^2)\log(1-r^2)}{2r^2} > 0$.

The positivity of the Lyapunov exponent $\gamma(e^{i\theta})$ implies (using the Ruelle–Osceledec theorem; see [30]) that there exists a constant $\lambda \neq 1$ (defining a boundary condition) for which

$$\lim_{n\to\infty} T_n(e^{i\theta}) \begin{pmatrix} 1 \\ \lambda \end{pmatrix} = 0. \tag{3.27}$$

From here we immediately get (using the theory of subordinate solutions developed in [30]) that for any $j$ and almost every $e^{i\theta} \in \partial\mathbb{D}$,

$$\lim_{k\to\infty} G_{j,j+k}(e^{i\theta}, \mathcal{C}) = 0. \tag{3.28}$$

We can use now (3.15) and (3.28) to verify the hypothesis of Lemma 3.2 for the random variables

$$X_k = \left| G_{j,j+k}(e^{i\theta}, \mathcal{C}) \right|^{s'}, \tag{3.29}$$

where $s' \in (s, 1)$. We therefore get

$$\lim_{k\to\infty} \mathbb{E}\left( \left| G_{j,j+k}\left(e^{i\theta}, \mathcal{C}\right) \right|^{s} \right) = 0. \qquad \square \tag{3.30}$$

The next step is to get the same result for the finite volume case (i.e., when we replace the matrix $\mathcal{C} = \mathcal{C}_\alpha$ by the matrix $\mathcal{C}_\alpha^{(n)}$).

**Lemma 3.4.** *For any fixed $j$, any $s \in (0, \frac{1}{2})$, and any $z \in \mathbb{D} \cup \partial\mathbb{D}$,*

$$\lim_{k\to\infty,\, k\leqslant n} \mathbb{E}\left( \left| G_{j,j+k}\left(z, \mathcal{C}_\alpha^{(n)}\right) \right|^{s} \right) = 0. \tag{3.31}$$

**Proof.** Let $\mathcal{C}$ be the CMV matrix corresponding to a family of Verblunsky coefficients $\{\alpha_n\}_{n\geqslant 0}$, with $|\alpha_n| < r$ for any $n$. Since $\mathbb{E}\left(\left|G_{j,j+k}(z,\mathcal{C})\right|^s\right) \to 0$ and $\mathbb{E}\left(\left|G_{j,j+k}(z,\mathcal{C})\right|^{2s}\right) \to 0$ as $k \to \infty$, we can take $k_\varepsilon \geqslant 0$ such that for any $k \geqslant k_\varepsilon$, $\mathbb{E}\left(\left|G_{j,j+k}(z,\mathcal{C})\right|^s\right) \leqslant \varepsilon$ and $\mathbb{E}\left(\left|G_{j,j+k}(z,\mathcal{C})\right|^{2s}\right) \leqslant \varepsilon$.

For $n \geqslant (k_\varepsilon + 2)$, let $\mathcal{C}^{(n)}$ be the CMV matrix obtained with the same $\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, \alpha_n, \ldots$ and with $\alpha_{n-1} \in \partial\mathbb{D}$. From now on we will use $G(z,\mathcal{C}) = (\mathcal{C} - z)^{-1}$ and $G(z, \mathcal{C}_\alpha^{(n)}) = (\mathcal{C}_\alpha^{(n)} - z)^{-1}$. Then

$$(\mathcal{C}_\alpha^{(n)} - z)^{-1} - (\mathcal{C} - z)^{-1} = (\mathcal{C} - z)^{-1}(\mathcal{C} - \mathcal{C}_\alpha^{(n)})(\mathcal{C}_\alpha^{(n)} - z)^{-1}. \tag{3.32}$$

Note that the matrix $(\mathcal{C}-\mathcal{C}^{(n)})$ has at most eight nonzero terms, each of absolute value at most 2. These nonzero terms are situated at positions $(m, m')$ and $|m-n| \leqslant 2$, $|m'-n| \leqslant 2$. Then

$$\mathbb{E}\left(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^s\right) \leqslant \mathbb{E}\left(|(\mathcal{C} - z)^{-1}_{j,j+k}|^s\right) + 2^s \sum_{8 \text{ terms}} \mathbb{E}\left(|(\mathcal{C} - z)^{-1}_{j,m}|^s \, |(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{m',j+k}|^s\right). \tag{3.33}$$

Using Schwarz's inequality,

$$\mathbb{E}\left(|\mathcal{C} - z)^{-1}_{j,m}|^s \, |(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{m',j+k}|^s\right) \leqslant \mathbb{E}\left(|\mathcal{C} - z)^{-1}_{j,m}|^{2s}\right)^{1/2} \mathbb{E}\left(|\mathcal{C}_\alpha^{(n)} - z)^{-1}_{m',j+k}|^{2s}\right)^{1/2}. \tag{3.34}$$

We clearly have $m \geqslant k_\varepsilon$ and therefore $\mathbb{E}\left(|\mathcal{C} - z)^{-1}_{j,m}|^{2s}\right) \leqslant \varepsilon$. Also, from Lemma 3.1, there exists a constant $C$ depending only on $s$ such that $\mathbb{E}\left(|\mathcal{C}_\alpha^{(n)} - z)^{-1}_{m',j+k}|^{2s}\right) \leqslant C$.

Therefore, for any $k \geqslant k_\varepsilon$, $\mathbb{E}\left(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^s\right) \leqslant \varepsilon + \varepsilon^{1/2} C$.

Since $\varepsilon$ is arbitrary, we obtain (3.31). $\square$

Note that Lemma 3.4 holds for any $s \in (0, \frac{1}{2})$. The result can be improved using a standard method:

**Lemma 3.5.** *For any fixed* $j$, *any* $s \in (0, 1)$, *and any* $z \in \mathbb{D}$,

$$\lim_{k \to \infty,\, k \leqslant n} \mathbb{E}\left(\left|G_{j,j+k}\left(z, \mathcal{C}_\alpha^{(n)}\right)\right|^s\right) = 0. \tag{3.35}$$

**Proof.** Let $s \in [\frac{1}{2}, 1)$, $t \in (s, 1)$, $r \in (0, \frac{1}{2})$. Then using the Hölder inequality for $p = \frac{t-r}{t-s}$ and for $q = \frac{t-r}{s-r}$, we get

$$\begin{aligned}
&\mathbb{E}\left(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^s\right) \\
&= \mathbb{E}\left(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^{\frac{r(t-s)}{t-r}} \, |(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^{\frac{t(s-r)}{t-r}}\right) \\
&\leqslant \left(\mathbb{E}\left(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^r\right)\right)^{\frac{t-s}{t-r}} \left(\mathbb{E}\left(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^t\right)\right)^{\frac{s-r}{t-r}}.
\end{aligned} \tag{3.36}$$

From Lemma 3.1, $\mathbb{E}(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^t)$ is bounded by a constant depending only on $t$ and from Lemma 3.4, $\mathbb{E}(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^r)$ tends to 0 as $k \to \infty$. We immediately get (3.35). $\square$

We can improve the previous lemma to get that the convergence to 0 of $\mathbb{E}(|(\mathcal{C}_\alpha^{(n)} - z)^{-1}_{j,j+k}|^s)$ is uniform in row $j$.

**Lemma 3.6.** *For any $\varepsilon > 0$, there exists a $k_\varepsilon \geqslant 0$ such that, for any $s, k, j, n, s \in (0,1), k > k_\varepsilon, n > 0, 0 \leqslant j \leqslant (n-1)$, and for any $z \in \mathbb{D} \cup \partial\mathbb{D}$, we have*

$$\mathbb{E}\left(\left|G_{j,j+k}\left(z, \mathcal{C}_\alpha^{(n)}\right)\right|^s\right) < \varepsilon. \tag{3.37}$$

**Proof.** As in the previous lemma, it is enough to prove the result for all $z \in \mathbb{D}$. Suppose the matrix $\mathcal{C}^{(n)}$ is obtained from the Verblunsky coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$. Let us consider the matrix $\mathcal{C}_{\text{dec}}^{(n)}$ obtained from the same Verblunsky coefficients with the additional restriction $\alpha_m = e^{i\theta}$ where $m$ is chosen to be bigger but close to $j$ (for example $m = j+3$). We will now compare $(\mathcal{C}^{(n)} - z)^{-1}_{j,j+k}$ and $(\mathcal{C}_{\text{dec}}^{(n)} - z)^{-1}_{j,j+k}$. By the resolvent identity,

$$\left|(\mathcal{C}^{(n)} - z)^{-1}_{j,j+k}\right| = \left|(\mathcal{C}^{(n)} - z)^{-1}_{j,j+k} - (\mathcal{C}_{\text{dec}}^{(n)} - z)^{-1}_{j,j+k}\right| \tag{3.38}$$

$$\leqslant 2 \sum_{|l-m| \leqslant 2, |l'-m| \leqslant 2} \left|(\mathcal{C}^{(n)} - z)^{-1}_{j,l}\right| \left|(\mathcal{C}_{\text{dec}}^{(n)} - z)^{-1}_{l',j+k}\right|. \tag{3.39}$$

The matrix $(\mathcal{C}_{\text{dec}}^{(n)} - z)^{-1}$ decouples between $m-1$ and $m$. Also, since $|l'-m| \leqslant 2$, we get that for any fixed $\varepsilon > 0$, we can pick a $k_\varepsilon$ such that for any $k \geqslant k_\varepsilon$ and any $l'$, $|l'-m| \leqslant 2$, we have

$$\mathbb{E}\left(\left|(\mathcal{C}_{\text{dec}}^{(n)} - z)^{-1}_{l',j+k}\right|\right) \leqslant \varepsilon. \tag{3.40}$$

(In other words, the decay is uniform on the 5 rows $m-2, m-1, m, m+1$, and $m+2$ situated at distance at most 2 from the place where the matrix $\mathcal{C}_{\text{dec}}^{(n)}$ decouples.)

As in Lemma 3.4, we can now use Schwarz's inequality to get that for any $\varepsilon > 0$ and for any $s \in (0, \frac{1}{2})$ there exists a $k_\varepsilon$ such that for any $j$ and any $k \geqslant k_\varepsilon$,

$$\mathbb{E}\left(\left|(\mathcal{C}^{(n)} - z)^{-1}_{j,j+k}\right|^s\right) < \varepsilon. \tag{3.41}$$

Using the same method as in Lemma 3.5, we get (3.37) for any $s \in (0,1)$. $\square$

We are heading towards proving the exponential decay of the fractional moments of the matrix elements of the resolvent of the CMV matrix. We will first prove a lemma about the behavior of the entries in the resolvent of the CMV matrix.

**Lemma 3.7.** *Suppose the random CMV matrix $\mathcal{C}^{(n)} = \mathcal{C}_\alpha^{(n)}$ is given as before (i.e., $\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, \alpha_{n-1}$ are independent random variables, the first $(n-1)$ uniformly distributed inside a disk of radius r and the last one uniformly distributed on the unit circle). Then, for any point $e^{i\theta} \in \partial\mathbb{D}$ and for any $\alpha \in \Omega$ where $G(e^{i\theta}, \mathcal{C}_\alpha^{(n)}) = (\mathcal{C}_\alpha^{(n)} - e^{i\theta})^{-1}$ exists, we have*

$$\frac{\left|G_{kl}(e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|}{\left|G_{ij}(e^{i\theta}, \mathcal{C}_\alpha^{(n)})\right|} \leqslant \left(\frac{2}{\sqrt{1-r^2}}\right)^{|k-i|+|l-j|}. \tag{3.42}$$

**Proof.** Using the results from Chapter 4 in Simon [29], the matrix elements of the resolvent of the CMV matrix are given by the following formulae:

$$\left[(\mathcal{C}-z)^{-1}\right]_{kl} = \begin{cases} (2z)^{-1}\chi_l(z)p_k(z), & k>l \quad \text{or} \quad k=l=2n-1, \\ (2z)^{-1}\pi_l(z)x_k(z), & l>k \quad \text{or} \quad k=l=2n, \end{cases} \tag{3.43}$$

where the polynomials $\chi_l(z)$ are obtained by the Gram–Schmidt process applied to $\{1, z, z^{-1}, \ldots\}$ in $L^2(\partial\mathbb{D}, d\mu)$ and the polynomials $x_k(z)$ are obtained by the Gram–Schmidt process applied to $\{1, z^{-1}, z \ldots\}$ in $L^2(\partial\mathbb{D}, d\mu)$. Also, $p_n$ and $\pi_n$ are the analogs of the Weyl solutions of Golinskii–Nevai [18] and are defined by

$$p_n = y_n + F(z)x_n, \tag{3.44}$$

$$\pi_n = \Upsilon_n + F(z)\chi_n, \tag{3.45}$$

where $y_n$ and $\Upsilon_n$ are the second kind analogs of the CMV bases and are given by

$$y_n = \begin{cases} z^{-l}\psi_{2l}, & n=2l, \\ -z^{-l}\psi^*_{2l-1}, & n=2l-1, \end{cases} \tag{3.46}$$

$$\Upsilon_n = \begin{cases} -z^{-l}\psi^*_{2l}, & n=2l, \\ z^{-l+1}\psi_{2l-1}, & n=2l-1, \end{cases} \tag{3.47}$$

The functions $\psi_n$ are the second kind polynomials associated to the measure $\mu$ and $F(z)$ is the Carathéodory function corresponding to $\mu$ (see [29]).

We will be interested in the values of the resolvent on the unit circle (we know they exist a.e. for the random matrices considered here). For any $z \in \partial\mathbb{D}$, the values of $F(z)$ are purely imaginary and also $\overline{\chi_n(z)} = x_n(z)$ and $\overline{\Upsilon_n(z)} = -y_n(z)$. In particular, $|\chi_n(z)| = |x_n(z)|$ for any $z \in \partial\mathbb{D}$.

Therefore $\overline{\pi_n(z)} = \overline{\Upsilon_n(z)} + \overline{F(z)\chi_n(z)} = -p_n(z)$, so $|\overline{\pi_n(z)}| = |p_n(z)|$ for any $z \in \partial\mathbb{D}$. We will also use $|\chi_{2n+1}(z)| = |\varphi_{2n+1}(z)|$, $|\chi_{2n}(z)| = |\varphi^*_{2n}(z)|$, $|x_{2n}(z)| = |\varphi_{2n}(z)|$, and $|x_{2n-1}(z)| = |\varphi^*_{2n-1}(z)|$ for any $z \in \partial\mathbb{D}$. Also, from Section 1.5 in [29], we have

$$\left|\frac{\varphi_{n\pm1}(z)}{\varphi_n(z)}\right| \leqslant C \tag{3.48}$$

for any $z \in \partial\mathbb{D}$, where $C = 2/\sqrt{1-r^2}$.

The key fact for proving (3.48) is that the orthonormal polynomials $\varphi_n$ satisfy a recurrence relation

$$\varphi_{n+1}(z) = \rho_n^{-1}(z\varphi_n(z) - \overline{\alpha}_n\varphi_n^*(z)). \tag{3.49}$$

This immediately gives the corresponding recurrence relation for the second kind polynomials

$$\psi_{n+1}(z) = \rho_n^{-1}(z\psi_n(z) + \overline{\alpha}_n\psi_n^*(z)). \tag{3.50}$$

Using (3.49) and (3.50), we will now prove a similar recurrence relation for the polynomials $\pi_n$. For any $z \in \partial\mathbb{D}$, we have

$$\begin{aligned}\pi_{2l+1}(z) &= \Upsilon_{2l+1}(z) + F(z)\chi_{2l+1}(z)\\ &= z^{-l}(\psi_{2l+1}(z) + F(z)\varphi_{2l+1}(z))\\ &= -\rho_{2l}^{-1}\, z\, \overline{\pi_{2l}(z)} + \rho_{2l}^{-1}\, \overline{\alpha}_{2l}\, \pi_{2l}(z)\end{aligned} \tag{3.51}$$

and similarly we get

$$\pi_{2l}(z) = -\rho_{2l-1}^{-1}\overline{\pi_{2l-1}(z)} - \alpha_{2l-1}\rho_{2l-1}^{-1}\pi_{2l-1}(z), \tag{3.52}$$

where we used the fact that for any $z \in \mathbb{D}$, $F(z)$ is purely imaginary, hence $\overline{F(z)} = -F(z)$.

Since $\rho_n^{-1} \leqslant \frac{1}{\sqrt{1-r^2}}$, Eqs. (3.51) and (3.52) will give that for any integer $n$ and any $z \in \mathbb{D}$,

$$\left|\frac{\pi_{n\pm1}(z)}{\pi_n(z)}\right| \leqslant C, \tag{3.53}$$

where $C = 2/\sqrt{1-r^2}$.

Using these observations and (3.43) we get, for any $z \in \partial\mathbb{D}$,

$$\left|\left[(\mathcal{C}^{(n)} - z)^{-1}\right]_{k,l}\right| \leqslant C\left|\left[(\mathcal{C}^{(n)} - z)^{-1}\right]_{k,l\pm1}\right| \tag{3.54}$$

and also

$$\left|\left[(\mathcal{C}^{(n)} - z)^{-1}\right]_{k,l}\right| \leqslant C\left|\left[(\mathcal{C}^{(n)} - z)^{-1}\right]_{k\pm1,l}\right|. \tag{3.55}$$

We can now combine (3.54) and (3.55) to get (3.42). □

We will now prove a simple lemma which will be useful in computations.

**Lemma 3.8.** *For any $s \in (0, 1)$ and any constant $\beta \in \mathbb{C}$, we have*

$$\int_{-1}^{1} \frac{1}{|x-\beta|^s}\, dx \leqslant \int_{-1}^{1} \frac{1}{|x|^s}\, dx. \tag{3.56}$$

**Proof.** Let $\beta = \beta_1 + i\beta_2$ with $\beta_1, \beta_2 \in \mathbb{R}$. Then

$$\int_{-1}^{1} \frac{1}{|x-\beta|^s}\, dx = \int_{-1}^{1} \frac{1}{|(x-\beta_1)^2 + \beta_2^2|^{s/2}}\, dx \leqslant \int_{-1}^{1} \frac{1}{|x-\beta_1|^s}\, dx. \tag{3.57}$$

But $1/|x|^s$ is the symmetric decreasing rearrangement of $1/|x-\beta_1|^s$ so we get

$$\int_{-1}^{1} \frac{1}{|x-\beta_1|^s}\, dx \leqslant \int_{-1}^{1} \frac{1}{|x|^s}\, dx \tag{3.58}$$

and therefore we immediately obtain (3.56). □

The following lemma shows that we can control conditional expectations of the diagonal elements of the matrix $\mathcal{C}^{(n)}$.

**Lemma 3.9.** *For any $s \in (0,1)$, any $k$, $1 \leqslant k \leqslant n$, and any choice of $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_{n-2}, \alpha_{n-1}$,*

$$\mathbb{E}\left( \left| F_{kk}(z, \mathcal{C}_\alpha^{(n)}) \right|^s \ \Big| \ \{\alpha_i\}_{i \neq k} \right) \leqslant C, \tag{3.59}$$

*where a possible value for the constant is $C = \frac{4}{1-s}\, 32^s$.*

**Proof.** For a fixed family of Verblunsky coefficients $\{\alpha_n\}_{n\in\mathbb{N}}$, the diagonal elements of the resolvent of the CMV matrix $\mathcal{C}$ can be obtained using the formula

$$(\delta_k, (\mathcal{C}+z)(\mathcal{C}-z)^{-1}\delta_k) = \int_{\partial\mathbb{D}} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, |\varphi_k(e^{i\theta})|^2\, d\mu(e^{i\theta}), \tag{3.60}$$

where $\mu$ is the measure on $\partial\mathbb{D}$ associated with the Verblunsky coefficients $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\varphi_n\}_{n\in\mathbb{N}}$ are the corresponding normalized orthogonal polynomials.

Using the results of Khrushchev [24], the Schur function of the measure $|\varphi_k(e^{i\theta})|^2\, d\mu(e^{i\theta})$ is

$$g_k(z) = f(z; \alpha_k, \alpha_{k+1}, \ldots)\, f(z; -\overline{\alpha}_{k-1}, -\overline{\alpha}_{k-2}, \ldots, -\overline{\alpha}_0, 1), \tag{3.61}$$

where by $f(z; S)$ we denote the Schur function associated to the family of Verblunsky coefficients $S$.

Since the dependence of $f(z; \alpha_k, \alpha_{k+1}, \ldots)$ on $\alpha_k$ is given by

$$f(z; \alpha_k, \alpha_{k+1}, \ldots) = \frac{\alpha_k + z f(z; \alpha_{k+1}, \alpha_{k+2} \ldots)}{1 + \overline{\alpha}_k z f(z; \alpha_{k+1}, \alpha_{k+2} \ldots)}, \tag{3.62}$$

we get that the dependence of $g_k(z)$ on $\alpha_k$ is given by

$$g_k(z) = C_1 \frac{\alpha_k + C_2}{1 + \overline{\alpha}_k C_2}, \tag{3.63}$$

where

$$C_1 = f(z; -\overline{\alpha}_{k-1}, -\overline{\alpha}_{k-2}, \ldots, -\overline{\alpha}_0, 1), \tag{3.64}$$

$$C_2 = z f(z; \alpha_{k+1}, \alpha_{k+2}, \ldots). \tag{3.65}$$

Note that the numbers $C_1$ and $C_2$ do not depend on $\alpha_k$, $|C_1|, |C_2| \leqslant 1$.

We now evaluate the Carathéodory function $F(z; |\varphi_k(e^{i\theta})|^2\, d\mu(e^{i\theta}))$ associated to the measure $|\varphi_k(e^{i\theta})|^2\, d\mu(e^{i\theta})$. By definition,

$$F(z; |\varphi_k(e^{i\theta})|^2\, d\mu(e^{i\theta})) = \int_{\partial\mathbb{D}} \frac{e^{i\theta}+z}{e^{i\theta}-z}\, |\varphi_k(e^{i\theta})|^2\, d\mu(e^{i\theta}) \tag{3.66}$$

$$= (\delta_k, (\mathcal{C}+z)(\mathcal{C}-z)^{-1}\delta_k). \tag{3.67}$$

We now have

$$\left| F(z; |\varphi_k(e^{i\theta})|^2\, d\mu(e^{i\theta})) \right| = \left| \frac{1 + z g_k(z)}{1 - z g_k(z)} \right| \leqslant \left| \frac{2}{1 - z\, C_1 \, \frac{\alpha_k + C_2}{1 + \overline{\alpha}_k C_2}} \right|. \tag{3.68}$$

It suffices to prove

$$\sup_{w_1,w_2\in\mathbb{D}} \int_{D(0,r)} \left| \frac{2}{1-w_1 \frac{\alpha_k+w_2}{1+\overline{\alpha}_k w_2}} \right|^s d\alpha_k \ < \ \infty. \tag{3.69}$$

Clearly

$$\begin{aligned} \left| \frac{2}{1-w_1 \frac{\alpha_k+w_2}{1+\overline{\alpha}_k w_2}} \right| &= \left| \frac{2(1+\overline{\alpha}_k w_2)}{1+\overline{\alpha}_k w_2 - w_1(\alpha_k+w_2)} \right| \\ &\leqslant \left| \frac{4}{1+\overline{\alpha}_k w_2 - w_1(\alpha_k+w_2)} \right|. \end{aligned} \tag{3.70}$$

For $\alpha_k = x+iy$, $1+\overline{\alpha}_k w_2 - w_1(\alpha_k+w_2) = x(-w_1+w_2)+y(-iw_1-iw_2)+(1-w_1w_2)$. Since for $w_1, w_2 \in \mathbb{D}$, $(-w_1+w_2)$, $(-iw_1-iw_2)$, and $(1-w_1w_2)$ cannot be all small, we will be able to prove (3.69).

If $|-w_1+w_2| \geqslant \varepsilon$,

$$\int_{D(0,r)} \left| \frac{2}{1-w_1 \frac{\alpha_k+w_2}{1+\overline{\alpha}_k w_2}} \right|^s d\alpha_k \leqslant \left(\frac{4}{\varepsilon}\right)^s \int_{-r}^{r}\int_{-r}^{r} \frac{1}{|x+yD+E|^s}\, dx\, dy \tag{3.71}$$

$$\leqslant 2\left(\frac{4}{\varepsilon}\right)^s \int_{-1}^{1} \frac{1}{|x|^s}\, dx = \frac{4}{1-s}\left(\frac{4}{\varepsilon}\right)^s \tag{3.72}$$

(where for the last inequality we used Lemma 3.8).

The same bound can be obtained for $|w_1+w_2| \geqslant \varepsilon$.

If $|-w_1+w_2| \leqslant \varepsilon$ and $|w_1+w_2| \leqslant \varepsilon$, then

$$|x(-w_1+w_2)+y(-iw_1-iw_2)+(1-w_1w_2)| \geqslant (1-\varepsilon^2-4\varepsilon) \tag{3.73}$$

so

$$\int_{D(0,r)} \left| \frac{2}{1-w_1 \frac{\alpha_k+w_2}{1+\overline{\alpha}_k w_2}} \right|^s d\alpha_k \leqslant 2^{s+2}\left(\frac{1}{1-\varepsilon^2-4\varepsilon}\right)^s. \tag{3.74}$$

Therefore for any small $\varepsilon$, we get (3.59) with

$$C = \max\left\{ \frac{4}{1-s}\left(\frac{4}{\varepsilon}\right)^s, 2^{s+2}\left(\frac{1}{1-\varepsilon^2-4\varepsilon}\right)^s \right\}. \tag{3.75}$$

For example, for $\varepsilon = \frac{1}{8}$, we get $C = \frac{4}{1-s}\, 32^s$. $\square$

We will now be able to prove Theorem 2.1.

**Proof of Theorem 2.1.** We will use the method developed by Aizenman et al. [3] for Schrödinger operators. The basic idea is to use the uniform decay of the expectations of the fractional moments of the matrix elements of $\mathcal{C}^{(n)}$ (Lemma 3.6) to derive the exponential decay.

We consider the matrix $\mathcal{C}^{(n)}$ obtained for the Verblunsky coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$. Fix a $k$, with $0 \leqslant k \leqslant (n-1)$. Let $\mathcal{C}_1^{(n)}$ be the matrix obtained for the Verblunsky coefficients

$\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ with the additional condition $\alpha_{k+m} = 1$ and $\mathcal{C}_2^{(n)}$ the matrix obtained from $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ with the additional restriction $\alpha_{k+m+3} = e^{i\theta}$ ($m$ is an integer $\geqslant 3$ which will be specified later, and $e^{i\theta}$ is a random point uniformly distributed on $\partial\mathbb{D}$).

Using the resolvent identity, we have

$$(\mathcal{C}^{(n)} - z)^{-1} - (\mathcal{C}_1^{(n)} - z)^{-1} = (\mathcal{C}_1^{(n)} - z)^{-1} (\mathcal{C}_1^{(n)} - \mathcal{C}^{(n)}) (\mathcal{C}^{(n)} - z)^{-1} \tag{3.76}$$

and

$$(\mathcal{C}^{(n)} - z)^{-1} - (\mathcal{C}_2^{(n)} - z)^{-1} = (\mathcal{C}^{(n)} - z)^{-1} (\mathcal{C}_2^{(n)} - \mathcal{C}^{(n)}) (\mathcal{C}_2^{(n)} - z)^{-1}. \tag{3.77}$$

Combining (3.76) and (3.77), we get

$$\begin{aligned}(\mathcal{C}^{(n)} - z)^{-1} = {} & (\mathcal{C}_1^{(n)} - z)^{-1} + (\mathcal{C}_1^{(n)} - z)^{-1} (\mathcal{C}_1^{(n)} - \mathcal{C}^{(n)}) (\mathcal{C}_2^{(n)} - z)^{-1} \\ & + (\mathcal{C}_1^{(n)} - z)^{-1} (\mathcal{C}_1^{(n)} - \mathcal{C}^{(n)}) (\mathcal{C}^{(n)} - z)^{-1} \\ & \times (\mathcal{C}_2^{(n)} - \mathcal{C}^{(n)}) (\mathcal{C}_2^{(n)} - z)^{-1},\end{aligned} \tag{3.78}$$

For any $k, l$ with $l \geqslant (k+m)$, we have

$$\left[(\mathcal{C}_1^{(n)} - z)^{-1}\right]_{kl} = 0 \tag{3.79}$$

and

$$\left[(\mathcal{C}_1^{(n)} - z)^{-1} (\mathcal{C}_1^{(n)} - \mathcal{C}^{(n)}) (\mathcal{C}_2^{(n)} - z)^{-1}\right]_{kl} = 0. \tag{3.80}$$

Therefore, since each of the matrices $(\mathcal{C}_1^{(n)} - \mathcal{C}^{(n)})$ and $(\mathcal{C}_2^{(n)} - \mathcal{C})$ has at most eight nonzero entries, we get that

$$\begin{aligned}\left[(\mathcal{C}^{(n)} - z)^{-1}\right]_{kl} = & \sum_{64 \text{ terms}} (\mathcal{C}_1^{(n)} - z)^{-1}_{ks_1} (\mathcal{C}_1^{(n)} - \mathcal{C}^{(n)})_{s_1 s_2} \\ & \times (\mathcal{C}^{(n)} - z)^{-1}_{s_2 s_3} (\mathcal{C}_2^{(n)} - \mathcal{C}^{(n)})_{s_3 s_4} (\mathcal{C}_2^{(n)} - z)^{-1}_{s_4 l}\end{aligned} \tag{3.81}$$

which gives

$$\begin{aligned}& \mathbb{E}\left(\left|(\mathcal{C}^{(n)} - z)^{-1}_{kl}\right|^s\right) \\ & \quad \leqslant 4^s \sum_{64 \text{ terms}} \mathbb{E}\left(\left|(\mathcal{C}_1^{(n)} - z)^{-1}_{ks_1} (\mathcal{C}^{(n)} - z)^{-1}_{s_2 s_3} (\mathcal{C}_2^{(n)} - z)^{-1}_{s_4 l}\right|^s\right),\end{aligned} \tag{3.82}$$

where since the matrix $\mathcal{C}_1^{(n)}$ decouples at $(k+m)$, we have $|s_2 - (k+m)| \leqslant 2$ and, since the matrix $\mathcal{C}_1^{(n)}$ decouples at $(k+m+3)$, we have $|s_3 - (k+m+3)| \leqslant 2$.

By Lemma 3.7, we have for any $e^{i\theta} \in \partial\mathbb{D}$,

$$\frac{\left|(\mathcal{C}^{(n)} - e^{i\theta})^{-1}_{s_2 s_3}\right|}{\left|(\mathcal{C}^{(n)} - e^{i\theta})^{-1}_{k+m+1,k+m+1}\right|} \leqslant \left(\frac{2}{\sqrt{1-r^2}}\right)^7. \tag{3.83}$$

Observe that $(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}$ and $(\mathcal{C}_2^{(n)}-z)^{-1}_{s_4l}$ do not depend on $\alpha_{k+m+1}$, and therefore using Lemma 3.9, we get

$$\begin{aligned}&\mathbb{E}\left(\left|(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}\,(\mathcal{C}^{(n)}-z)^{-1}_{s_2s_3}\,(\mathcal{C}_2^{(n)}-z)^{-1}_{s_4l}\right|^s\ \Big|\ \{\alpha_i\}_{i\neq(k+m+1)}\right)\\ &\quad\leqslant\frac{4}{1-s}\,32^s\left(\frac{2}{\sqrt{1-r^2}}\right)^7\left|(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}\right|^s\left|(\mathcal{C}_2^{(n)}-z)^{-1}_{s_4l}\right|^s.\end{aligned} \tag{3.84}$$

Since the random variables $(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}$ and $(\mathcal{C}_2^{(n)}-z)^{-1}_{s_4l}$ are independent (they depend on different sets of Verblunsky coefficients), we get

$$\begin{aligned}&\mathbb{E}\left(\left|(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}\,(\mathcal{C}^{(n)}-z)^{-1}_{s_2s_3}\,(\mathcal{C}_2^{(n)}-z)^{-1}_{s_4l}\right|^s\right)\\ &\quad\leqslant C(s,r)\,\mathbb{E}\left(\left|(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}\right|^s\right)\mathbb{E}\left(\left|(\mathcal{C}_1^{(n)}-z)^{-1}_{s_4l}\right|^s\right),\end{aligned} \tag{3.85}$$

where $C(s,r)=\frac{4}{1-s}\,32^s\left(\frac{2}{\sqrt{1-r^2}}\right)^7$.

The idea for obtaining exponential decay is to use the terms $\mathbb{E}(|(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}|^s)$ to get smallness and the terms $\mathbb{E}(|(\mathcal{C}_1^{(n)}-z)^{-1}_{s_4l}|^s)$ to repeat the process. Thus, using the Lemma 3.6, we get that for any $\beta<1$, there exists a fixed constant $m\geqslant 0$ such that, for any $s_1$, $|s_1-(k+m)|\leqslant 2$, we have

$$4^s\cdot 64\cdot C(s,r)\cdot\mathbb{E}\left(\left|(\mathcal{C}_1^{(n)}-z)^{-1}_{ks_1}\right|^s\right)<\beta. \tag{3.86}$$

We can now repeat the same procedure for each term $\mathbb{E}(|(\mathcal{C}_1^{(n)}-z)^{-1}_{s_4l}|^s)$ and we gain one more coefficient $\beta$. At each step, we move $(m+3)$ spots to the right from $k$ to $l$. We can repeat this procedure $\left[\frac{l-k}{m+3}\right]$ times and we get

$$\mathbb{E}\left(\left|(\mathcal{C}^{(n)}-z)^{-1}_{kl}\right|^s\right)\leqslant C\beta^{(l-k)/(m+3)} \tag{3.87}$$

which immediately gives (2.6). $\square$

## 4. The localized structure of the eigenfunctions

In this section, we will study the eigenfunctions of the random CMV matrices considered in (2.1). We will prove that, with probability 1, each eigenfunction of these matrices will be exponentially localized about a certain point, called the center of localization. We will follow ideas from del Rio et al. [8].

Theorem 2.1 will give that, for any $z\in\partial\mathbb{D}$, any integer $n$ and any $s\in(0,1)$,

$$\mathbb{E}\left(\left|F_{kl}(z,\mathcal{C}_\alpha^{(n)})\right|^s\right)\leqslant C\,e^{-D|k-l|}. \tag{4.1}$$

Aizenman's theorem for CMV matrices (see [31]) shows that (4.1) implies that for some positive constants $C_0$ and $D_0$ depending on $s$, we have

$$\mathbb{E}\left(\sup_{j\in\mathbb{Z}} \left|\left(\delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l\right)\right|\right) \leqslant C_0\, e^{-D_0|k-l|}. \tag{4.2}$$

This will allow us to conclude that the eigenfunctions of the CMV matrix are exponentially localized. The first step will be

**Lemma 4.1.** *For almost every $\alpha \in \Omega$, there exists a constant $D_\alpha > 0$ such that for any $n$, any $k, l$, with $1 \leqslant k, l \leqslant n$, we have*

$$\sup_{j\in\mathbb{Z}} |(\delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l)| \leqslant D_\alpha\, (1+n)^6\, e^{-D_0|k-l|}. \tag{4.3}$$

**Proof.** From (4.2) we get that

$$\int_\Omega \left(\sup_{j\in\mathbb{Z}} \left|\left(\delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l\right)\right|\right) d\mathbb{P}(\alpha) \leqslant C_0\, e^{-D_0|k-l|} \tag{4.4}$$

and therefore there exists a constant $C_1 > 0$ such that

$$\begin{aligned}\sum_{n,k,l=1\ k,l\leqslant n}^{\infty} \int_\Omega & \frac{1}{(1+n)^2(1+k)^2(1+l)^2} \\ & \times \left(\sup_{j\in\mathbb{Z}} \left|\left(\delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l\right)\right|\right) e^{D_0|k-l|}\, d\mathbb{P}(\alpha) \leqslant C_1.\end{aligned} \tag{4.5}$$

It is clear that for any $k, l$, with $1 \leqslant k, l \leqslant n$, the function

$$\alpha \longrightarrow \frac{1}{(1+n)^2(1+k)^2(1+l)^2} \left(\sup_{j\in\mathbb{Z}} \left|\left(\delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l\right)\right|\right) e^{D_0|k-l|} \tag{4.6}$$

is integrable.

Hence, for almost every $\alpha \in \Omega$, there exists a constant $D_\alpha > 0$ such that for any $n, k, l$, with $1 \leqslant k, l \leqslant n$,

$$\sup_{j\in\mathbb{Z}} \left|\left(\delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l\right)\right| \leqslant D_\alpha\, (1+n)^6\, e^{-D_0|k-l|}. \qquad \square \tag{4.7}$$

A useful version of the previous lemma is

**Lemma 4.2.** *For almost every $\alpha \in \Omega$, there exists a constant $C_\alpha > 0$ such that for any $n$, any $k, l$, with $1 \leqslant k, l \leqslant n$, and $|k-l| \geqslant \frac{12}{D_0} \ln(n+1)$, we have*

$$\sup_{j\in\mathbb{Z}} |(\delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l)| \leqslant C_\alpha\, e^{-\frac{D_0}{2}|k-l|}. \tag{4.8}$$

**Proof.** It is clear that for any $n, k, l$, with $1 \leqslant k, l \leqslant n$ and $|k-l| \geqslant \frac{12}{D_0} \ln(n+1)$,

$$\frac{1}{(1+n^2)(1+k^2)(1+l^2)}\, e^{\frac{D_0}{2}|k-l|} \geqslant 1. \tag{4.9}$$

In particular, for any $n, k, l$ with $|k-l| \geqslant \frac{12}{D_0} \ln(n+1)$, the function

$$\Omega \ni \alpha \longrightarrow \left( \sup_{j \in \mathbb{Z}} \left| \left( \delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l \right) \right| \right) e^{\frac{D_0}{2}|k-l|} \tag{4.10}$$

is integrable, so it is finite for almost every $\alpha$.

Hence for almost every $\alpha \in \Omega$, there exists a constant $C_\alpha > 0$ such that for any $k, l$, $|k-l| \geqslant \frac{12}{D_0} \ln(n+1)$,

$$\sup_{j \in \mathbb{Z}} \left| \left( \delta_k, (\mathcal{C}_\alpha^{(n)})^j \delta_l \right) \right| \leqslant C_\alpha\, e^{-\frac{D_0}{2}|k-l|}. \qquad \square \tag{4.11}$$

**Proof of Theorem 2.2.** Let us start with a CMV matrix $\mathcal{C}^{(n)} = \mathcal{C}_\alpha^{(n)}$ corresponding to the Verblunsky coefficients $\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, \alpha_{n-1}$. As mentioned before, the spectrum of $\mathcal{C}^{(n)}$ is simple. Let $e^{i\theta_\alpha}$ be an eigenvalue of the matrix $\mathcal{C}_\alpha^{(n)}$ and $\varphi_\alpha^{(n)}$ a corresponding eigenfunction.

We see that, on the unit circle, the sequence of functions

$$f_M(e^{i\theta}) = \frac{1}{2M+1} \sum_{j=-M}^{M} e^{ij(\theta-\theta_\alpha)} \tag{4.12}$$

is uniformly bounded (by 1) and converge pointwise (as $M \to \infty$) to the characteristic function of the point $e^{i\theta_\alpha}$. Let $P_{\{e^{i\theta_\alpha}\}} = \chi_{\{e^{i\theta_\alpha}\}}(\mathcal{C}_\alpha^{(n)})$.

By Lemma 4.2, we have, for any $k, l$, with $|k-l| \geqslant \frac{12}{D_0} \ln(n+1)$,

$$\left| \left( \delta_k, f_M(\mathcal{C}_\alpha^{(n)})\, \delta_l \right) \right| = \frac{1}{2M+1} \left| \sum_{j=-M}^{M} \left( \delta_k, e^{-ij\,\theta_\alpha} (\mathcal{C}_\alpha^{(n)})^j\, \delta_l \right) \right| \tag{4.13}$$

$$\leqslant \frac{1}{2M+1} \sum_{j=-M}^{M} \left| \left( \delta_k, (\mathcal{C}_\alpha^{(n)})^j\, \delta_l \right) \right| \leqslant C_\alpha\, e^{-\frac{D_0}{2}|k-l|}, \tag{4.14}$$

where for the last inequality we used (4.1).

By taking $M \to \infty$ in the previous inequality, we get

$$\left| \left( \delta_k, P_{\{e^{i\theta_\alpha}\}}\, \delta_l \right) \right| \leqslant C_\alpha\, e^{-\frac{D_0}{2}|k-l|} \tag{4.15}$$

and therefore

$$\left| \varphi_\alpha^{(n)}(k)\, \varphi_\alpha^{(n)}(l) \right| \leqslant C_\alpha\, e^{-\frac{D_0}{2}|k-l|}. \tag{4.16}$$

We can now pick as the center of localization the smallest integer $m(\varphi_\alpha^{(n)})$ such that

$$|\varphi_\alpha^{(n)}(m(\varphi_\alpha^{(n)}))| = \max_m |\varphi_\alpha^{(n)}(m)|. \tag{4.17}$$

We clearly have $|\varphi_\alpha^{(n)}(m(\varphi_\alpha^{(n)}))| \geqslant \frac{1}{\sqrt{n+1}}$.

Using the inequality (4.16) with $k = m$ and $l = m(\varphi_\alpha^{(n)})$ we get, for any $m$ with $|m - m(\varphi_\alpha^{(n)})| \geqslant \frac{12}{D_0} \ln(n+1)$,

$$\left|\varphi_\alpha^{(n)}(m)\right| \leqslant C_\alpha\, e^{-\frac{D_0}{2}|m-m(\varphi_\alpha^{(n)})|}\sqrt{n+1}. \tag{4.18}$$

Since for large $n$, $e^{-\frac{D_0}{2}|k-l|}\sqrt{n+1} \leqslant e^{-\frac{D_0}{3}|k-l|}$ for any $k, l$, $|k-l| \geqslant \frac{12}{D_0}\ln(n+1)$, we get the desired conclusion (we can take $D_2 = \frac{12}{D_0}$). $\square$

For any eigenfunction $\varphi_\alpha^{(n)}$, the point $m(\varphi_\alpha^{(n)})$ is called its center of localization. The eigenfunction is concentrated (has its large values) near the point $m(\varphi_\alpha^{(n)})$ and is tiny at sites that are far from $m(\varphi_\alpha^{(n)})$. This structure of the eigenfunctions will allow us to prove a decoupling property of the CMV matrix.

Note that we used Lemma 4.2 in the proof of Theorem 2.2. We can get a stronger result by using Lemma 4.1 (we replace (4.11) by (4.7)). Thus, for any $n$ and any $m \leqslant n$, we have

$$\left|\varphi_\alpha^{(n)}(m)\right| \leqslant D_\alpha\,(1+n)^6\, e^{-\frac{D_0}{2}\,|m-m(\varphi_\alpha^{(n)})|}\sqrt{n+1}, \tag{4.19}$$

where $m(\varphi_\alpha^{(n)})$ is the center of localization of the eigenfunction $\varphi_\alpha^{(n)}$.

## 5. Decoupling the point process

We will now show that the distribution of the eigenvalues of the CMV matrix $\mathcal{C}^{(n)}$ can be approximated (as $n \to \infty$) by the distribution of the eigenvalues of another matrix CMV matrix $\tilde{\mathcal{C}}^{(n)}$, which decouples into the direct sum of smaller matrices.

As explained in Section 1, for the CMV matrix $\mathcal{C}^{(n)}$ obtained with the Verblunsky coefficients $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{n-1}) \in \Omega$, we consider $\tilde{\mathcal{C}}^{(n)}$ the CMV matrix obtained from the same Verblunsky coefficients with the additional restrictions $\alpha_{\left[\frac{n}{\ln n}\right]} = e^{i\eta_1}$, $\alpha_{2\left[\frac{n}{\ln n}\right]} = e^{i\eta_2}, \ldots, \alpha_{n-1} = e^{i\eta_{[\ln n]}}$, where $e^{i\eta_1}, e^{i\eta_2}, \ldots, e^{i\eta_{[\ln n]}}$ are independent random points uniformly distributed on the unit circle. The matrix $\tilde{\mathcal{C}}^{(n)}$ decouples into the direct sum of approximately $[\ln n]$ unitary matrices $\tilde{\mathcal{C}}_1^{(n)}, \tilde{\mathcal{C}}_2^{(n)}, \ldots, \tilde{\mathcal{C}}_{[\ln n]}^{(n)}$. Since we are interested in the asymptotic distribution of the eigenvalues, it will be enough to study the distribution (as $n \to \infty$) of the eigenvalues of the matrices $\mathcal{C}^{(N)}$ of size $N = [\ln n]\left[\frac{n}{\ln n}\right]$. Note that in this situation the corresponding truncated matrix $\tilde{\mathcal{C}}^{(N)}$ will decouple into the direct sum of exactly $[\ln n]$ identical blocks of size $\left[\frac{n}{\ln n}\right]$.

We will begin by comparing the matrices $\mathcal{C}^{(N)}$ and $\tilde{\mathcal{C}}^{(N)}$.

**Lemma 5.1.** *For* $N = [\ln n]\left[\frac{n}{\ln n}\right]$, *the matrix* $\mathcal{C}^{(N)} - \tilde{\mathcal{C}}^{(N)}$ *has at most* $4[\ln n]$ *nonzero rows.*

**Proof.** In our analysis, we will start counting the rows of the CMV matrix with row 0. A simple inspection of the CMV matrix shows that for even Verblunsky coefficients $\alpha_{2k}$, only the rows $2k$ and $2k+1$ depend on $\alpha_{2k}$. For odd Verblunsky coefficients $\alpha_{2k+1}$, only the rows $2k, 2k+1, 2k+2, 2k+3$ depend on $\alpha_{2k+1}$.

Since in order to obtain the matrix $\tilde{\mathcal{C}}^{(N)}$ from $\mathcal{C}^{(N)}$ we modify $[\ln n]$ Verblunsky coefficients $\alpha_{\left[\frac{n}{\ln n}\right]}, \alpha_{2\left[\frac{n}{\ln n}\right]}, \ldots, \alpha_{[\ln n]\left[\frac{n}{\ln n}\right]}$, we immediately see that at most $4[\ln n]$ rows of $\mathcal{C}^{(N)}$ are modified.

Therefore $\mathcal{C}^{(N)} - \tilde{\mathcal{C}}^{(N)}$ has at most $4[\ln n]$ nonzero rows (and, by the same argument, at most 4 columns around each point where the matrix $\tilde{\mathcal{C}}^{(N)}$ decouples). $\square$

Since we are interested in the points situated near the places where the matrix $\tilde{\mathcal{C}}^{(N)}$ decouples, a useful notation will be

$$S_N(K) = S^{(1)}(K) \cup S^{(2)}(K) \cup \cdots \cup S^{([\ln n])}(K), \tag{5.1}$$

where $S^{(k)}(K)$ is a set of $K$ integers centered at $k\left[\frac{n}{\ln n}\right]$ (e.g., for $K = 2p$, $S^{(k)}(K) = \left\{k\left[\frac{n}{\ln n}\right] - p + 1, k\left[\frac{n}{\ln n}\right] - p + 2, \ldots k\left[\frac{n}{\ln n}\right] + p\right\}$). Using this notation, we also have

$$S_N(1) = \left\{\left[\frac{n}{\ln n}\right], 2\left[\frac{n}{\ln n}\right], \ldots, [\ln n]\left[\frac{n}{\ln n}\right]\right\}. \tag{5.2}$$

Consider the intervals $I_{N,k}$, $1 \leqslant k \leqslant m$, of size $\frac{1}{N}$ near the point $e^{i\alpha}$ on the unit circle (for example $I_{N,k} = (e^{i(\alpha + \frac{a_k}{N})}, e^{i(\alpha + \frac{b_k}{N})})$), where $a_1 < b_1 \leqslant a_2 < b_2 \leqslant \cdots \leqslant a_m < b_m$. We will denote by $\mathcal{N}_N(I)$ the number of eigenvalues of $\mathcal{C}^{(N)}$ situated in the interval $I$, and by $\tilde{\mathcal{N}}_N(I)$ the number of eigenvalues of $\tilde{\mathcal{C}}^{(N)}$ situated in $I$. We will prove that, for large $N$, $\mathcal{N}_N(I_{N,k})$ can be approximated by $\tilde{\mathcal{N}}_N(I_{N,k})$, that is, for any integers $k_1, k_2, \ldots, k_m \geqslant 0$, we have, for $N \to \infty$,

$$\begin{aligned} \Big| \, &\mathbb{P}(\mathcal{N}_N(I_{N,1}) = k_1, \, \mathcal{N}_N(I_{N,2}) = k_2, \ldots, \mathcal{N}_N(I_{N,m}) = k_m) \\ &- \mathbb{P}(\tilde{\mathcal{N}}_N(I_{N,1}) = k_1, \, \tilde{\mathcal{N}}_N(I_{N,2}) = k_2, \ldots, \tilde{\mathcal{N}}_N(I_{N,m}) = k_m) \Big| \longrightarrow 0. \end{aligned} \tag{5.3}$$

Since, by the results in Section 4, the eigenfunctions of the matrix $\mathcal{C}^{(N)}$ are exponentially localized (supported on a set of size $2T[\ln(n+1)]$, where, from now on, $T = \frac{14}{D_0}$), some of them will have the center of localization near $S_N(1)$ (the set of points where the matrix $\tilde{\mathcal{C}}^{(N)}$ decouples) and others will have centers of localization away from this set (i.e., because of exponential localization, inside an interval $\left(k\left[\frac{n}{\ln n}\right], (k+1)\left[\frac{n}{\ln n}\right]\right)$).

Roughly speaking, each eigenfunction of the second type will produce an "almost" eigenfunction for one of the blocks of the decoupled matrix $\tilde{\mathcal{C}}^{(N)}$. These eigenfunctions will allow us to compare $\mathcal{N}_N(I_{N,k})$ and $\tilde{\mathcal{N}}_N(I_{N,k})$.

We see that any eigenfunction with the center of localization outside the set $S_N(4T[\ln n])$ will be tiny on the set $S_N(1)$. Therefore, if we want to estimate the number of eigenfunctions that are supported close to $S_N(1)$, it will be enough to analyze the number $b_{N,\alpha}$, where $b_{N,\alpha}$

= number of eigenfunctions of $\mathcal{C}_\alpha^{(N)}$ with the center of localization inside $S_N(4T[\ln n])$ (we will call these eigenfunctions "bad eigenfunctions"). We will now prove that the number $b_{N,\alpha}$ is small compared to $N$.

A technical complication is generated by the fact that in the exponential localization of eigenfunctions given by (4.3), the constant $D_\alpha$ depends on $\alpha \in \Omega$. We define

$$\mathcal{M}_K = \left\{ \alpha \in \Omega,\ \sup_{j \in \mathbb{Z}} |(\delta_k, (\mathcal{C}^{(N)})^j \delta_l)| \leqslant K\,(1+N)^6\, e^{-D_0|k-l|} \right\}. \tag{5.4}$$

Note that for any $K > 0$, the set $\mathcal{M}_K \subset \Omega$ is invariant under rotation. Also, we can immediately see that the sets $\mathcal{M}_K$ grow with $K$ and

$$\lim_{K \to \infty} \mathbb{P}(\mathcal{M}_K) = 1 \tag{5.5}$$

We will be able to control the number of "bad eigenfunctions" for $\alpha \in \mathcal{M}_K$ using the following lemma:

**Lemma 5.2.** *For any $K > 0$ and any $\alpha \in \mathcal{M}_K$, there exists a constant $C_K > 0$ such that*

$$b_{N,\alpha} \leqslant C_K\,(\ln(1+N))^2\,. \tag{5.6}$$

**Proof.** For any $K > 0$, any $\alpha \in \mathcal{M}_K$, and any eigenfunction $\varphi_\alpha^N$ which is exponentially localized about a point $m(\varphi_\alpha^N)$, we have, using (4.19),

$$\left|\varphi_\alpha^{(N)}(m)\right| \leqslant K\, e^{-\frac{D_0}{2}\,|m-m(\varphi_\alpha^{(N)})|}\,(1+N)^6\sqrt{1+N}. \tag{5.7}$$

Therefore for any $m$ such that $|m-m(\varphi_\alpha^N)| \geqslant \left[\frac{14}{D_0}\ln(1+N)\right]$, we have

$$\begin{aligned}\sum_{|m-m(\varphi_\alpha^N)| \geqslant \left[\frac{14}{D_0}\ln(1+N)\right]} \left|\varphi_\alpha^{(N)}(m)\right|^2 &\leqslant 2(1+N)^{-14}\,(1+N)^{13} \sum_{k=0}^{\infty} K^2\, e^{-D_0 k} \\ &\leqslant (1+N)^{-1} K^2 \frac{2e^{D_0}}{e^{D_0}-1}.\end{aligned} \tag{5.8}$$

Therefore, for any fixed $K$ and $s$, we can find an $N_0 = N_0(k,s)$ such that for any $N \geqslant N_0$,

$$\sum_{|m-m(\varphi_\alpha^N)| \leqslant \left[\frac{14}{D_0}\ln(1+N)\right]} \left|\varphi_\alpha^{(N)}(m)\right|^2 \geqslant \tfrac{1}{2}. \tag{5.9}$$

We will consider eigenfunctions $\varphi_\alpha^N$ with the center of localization in $S_N\,(4T[\ln N])$. For a fixed $\alpha \in \mathcal{M}_K$, we denote the number of these eigenfunctions by $b_{N,\alpha}$. We denote by $\{\psi_1, \psi_2, \ldots, \psi_{b_{N,\alpha}}\}$ the set of these eigenfunctions. Since the spectrum of $\mathcal{C}^{(N)}$ is simple, this is an orthonormal set.

Therefore, if we denote by card$(A)$ the number of elements of the set $A$, we get

$$\begin{aligned}
&\sum_{m\in S\left(4T[\ln N]+\left[\frac{14}{D_0}\ln(1+N)\right]\right)} \sum_{i=1}^{b_{N,\alpha}} |\psi_i(m)|^2 \\
&\quad\leqslant \text{card}\left\{S\left(4T[\ln N]+\left[\frac{14}{D_0}\ln(1+N)\right]\right)\right\} \\
&\quad\leqslant \left(4T+\frac{14}{D_0}\right)(\ln(1+N))^2 .
\end{aligned}$$

Also, from (5.9), for any $N \geqslant N_0(K,s)$,

$$\sum_{m\in S\left(4T[\ln N]+\left[\frac{14}{D_0}\ln(1+N)\right]\right)} \sum_{i=1}^{b_{N,\alpha}} |\psi_i(m)|^2 \geqslant \tfrac{1}{2}\, b_{N,\alpha}. \tag{5.10}$$

Therefore, for any $K > 0$ and any $\alpha \in \mathcal{M}_K$, we have, for $N \geqslant N_0(K,s)$,

$$b_{N,\alpha} \leqslant 2\left(4T+\frac{14}{D_0}\right)(\ln(1+N))^2 \tag{5.11}$$

and we can now conclude (5.6). □

Lemma 5.2 shows that for any $K \geqslant 0$, the number of "bad eigenfunctions" corresponding to $\alpha \in \mathcal{M}_K$ is of the order $(\ln N)^2$ (hence small compared to $N$).

Since the distributions for our Verblunsky coefficients are taken to be rotationally invariant, the distribution of the eigenvalues is rotationally invariant. Therefore, for any interval $I_N$ of size $\frac{1}{N}$ on the unit circle, and for any fixed set $\mathcal{M}_K \subset \Omega$, the expected number of "bad eigenfunctions" corresponding to eigenvalues in $I_N$ is of size $\frac{(\ln N)^2}{N}$. We then get that the probability of the event "there are bad eigenfunctions corresponding to eigenvalues in the interval $I_N$" converges to 0. This fact will allow us to prove

**Lemma 5.3.** *For any $K > 0$, any disjoint intervals $I_{N,1}, I_{N,2}, \ldots, I_{N,m}$ (each one of size $\frac{1}{N}$ and situated near the point $e^{i\alpha}$) and any positive integers $k_1, k_2, \ldots, k_m$, we have*

$$\begin{aligned}
\Big|\, &\mathbb{P}(\{\mathcal{N}_N(I_{N,1}) = k_1,\ \mathcal{N}_N(I_{N,2}) = k_2, \ldots, \mathcal{N}_N(I_{N,m}) = k_m\} \cap \mathcal{M}_K) \\
&\qquad - \mathbb{P}(\{\tilde{\mathcal{N}}_N(I_{N,1}) = k_1, \tilde{\mathcal{N}}_N(I_{N,2}) = k_2, \ldots, \tilde{\mathcal{N}}_N(I_{N,m}) = k_m\} \\
&\qquad \cap \mathcal{M}_K)\Big| \longrightarrow 0
\end{aligned} \tag{5.12}$$

*as $N \to \infty$.*

**Proof.** We will work with $\alpha \in \mathcal{M}_K$. We first observe that any "good eigenfunction" (i.e., an eigenfunction with the center of localization outside $S_N(4T[\ln N])$) is tiny on $S_N(1)$.

Indeed, from (4.19), for any eigenfunction $\varphi_\alpha^{(N)}$ with the center of localization $m(\varphi_\alpha^{(N)})$ and for any $m$ with $|m - m(\varphi_\alpha^{(N)})| \geqslant \frac{18}{D_0}[\ln(N+1)]$,

$$|\varphi_\alpha^{(N)}(m)| \leqslant K e^{-\frac{D_0}{2}|m-m(\varphi_\alpha^{(N)})|}(1+N)^6\sqrt{1+N}. \tag{5.13}$$

In particular, if the center of localization of $\varphi_\alpha^{(N)}$ is outside $S_N\,(4T[\ln N])$, then for all $m \in S_N(1)$, we have

$$|\varphi_\alpha^{(N)}(m)| \leqslant K(1+N)^{-2}. \tag{5.14}$$

We will use the fact that if $N$ is a normal matrix, $z_0 \in \mathbb{C}$, $\varepsilon > 0$, and $\varphi$ is a unit vector with

$$\| (N - z_0)\varphi \| < \varepsilon \tag{5.15}$$

then $N$ has an eigenvalue in $\{z \mid |z - z_0| < \varepsilon\}$.

For any "good eigenfunction" $\varphi_\alpha^{(N)}$, we have $\mathcal{C}_\alpha^{(N)}\varphi_\alpha^{(N)} = 0$ and therefore, using Lemma 5.1,

$$\|\tilde{\mathcal{C}}_\alpha^{(N)}\varphi_\alpha^{(N)}\| \leqslant 2K[\ln N](1+N)^{-2}. \tag{5.16}$$

Therefore, for any interval $I_N$ of size $\frac{1}{N}$, we have

$$\mathcal{N}_N(I_N) \leqslant \tilde{\mathcal{N}}_N(\tilde{I}_N), \tag{5.17}$$

where $\tilde{I}_N$ is the interval $I_N$ augmented by $2K[\ln N](1+N)^{-2}$.

Since $2K[\ln N](1+N)^{-2} = o(\frac{1}{N})$, we can now conclude that

$$\mathbb{P}\big(\big(\mathcal{N}_N(I_N) \leqslant \tilde{\mathcal{N}}_N(I_N)\big) \cap \mathcal{M}_K\big) \to 1 \quad \text{as} \quad n \to \infty. \tag{5.18}$$

We can use the same argument (starting from the eigenfunctions of $\tilde{\mathcal{C}}_\alpha^{(N)}$, which are also exponentially localized) to show that

$$\mathbb{P}\big(\big(\mathcal{N}_N(I_N) \geqslant \tilde{\mathcal{N}}_N(I_N)\big) \cap \mathcal{M}_K\big) \to 1 \quad \text{as} \quad n \to \infty, \tag{5.19}$$

so we can now conclude that

$$\mathbb{P}\big(\big(\mathcal{N}_N(I_N) = \tilde{\mathcal{N}}_N(I_N)\big) \cap \mathcal{M}_K\big) \to 1 \quad \text{as} \quad n \to \infty. \tag{5.20}$$

Instead of one interval $I_N$, we can take $m$ intervals $I_{N,1}, I_{N,2}, \ldots, I_{N,m}$ so we get (5.12). □

**Proof of Theorem 2.3.** Lemma 5.3 shows that for any $K > 0$, the distribution of the eigenvalues of the matrix $\mathcal{C}^{(N)}$ can be approximated by the distribution of the eigenvalues of the matrix $\tilde{\mathcal{C}}^{(N)}$ when we restrict to the set $\mathcal{M}_K \subset \Omega$. Since by (5.5) the sets $\mathcal{M}_K$ grow with $K$ and $\lim_{K\to\infty} \mathbb{P}(\mathcal{M}_K) = 1$, we get the desired result. □

## 6. Estimating the probability of having two or more eigenvalues in an interval

The results from the previous section show that the local distribution of the eigenvalues of the matrix $\mathcal{C}^{(N)}$ can be approximated by the direct sum of the local distribution of $[\ln n]$ matrices of size $\left[\frac{n}{\ln n}\right]$, $\mathcal{C}_1^{(N)}, \mathcal{C}_2^{(N)}, \ldots, \mathcal{C}_{[\ln n]}^{(N)}$. These matrices are decoupled and depend on independent sets of Verblunsky coefficients; hence they are independent.

For a fixed point $e^{i\theta_0} \in \partial\mathbb{D}$, and an interval $I_N = (e^{i(\theta_0+\frac{2\pi a}{N})}, e^{i(\theta_0+\frac{2\pi b}{N})})$, we will now want to control the probability of the event "$\mathcal{C}^{(N)}$ has $k$ eigenvalues in $I_N$." We will analyze the distribution of the eigenvalues of the direct sum of the matrices $\mathcal{C}_1^{(N)}, \mathcal{C}_2^{(N)}, \ldots, \mathcal{C}_{[\ln n]}^{(N)}$. We will prove that, as $n \to \infty$, each of the decoupled matrices $\mathcal{C}_k^{(N)}$ contributes (up to a negligible error) at most one eigenvalue in the interval $I_N$.

For any nonnegative integer $m$, denote by $A(m, \mathcal{C}, I)$ the event

$$A(m, \mathcal{C}, I) = \text{"}\mathcal{C} \text{ has at least } m \text{ eigenvalues in the interval } I\text{"} \tag{6.1}$$

and by $B(m, \mathcal{C}, I)$ the event

$$B(m, \mathcal{C}, I) = \text{"}\mathcal{C} \text{ has exactly } m \text{ eigenvalues in the interval } I\text{"} \tag{6.2}$$

In order to simplify future notations, for any point $e^{i\theta} \in \partial\mathbb{D}$, we also define the event $M(e^{i\theta})$ to be

$$M(e^{i\theta}) = \text{"}e^{i\theta} \text{ is an eigenvalue of } \mathcal{C}^{(N)}\text{"} \tag{6.3}$$

We can begin by observing that the eigenvalues of the matrix $\mathcal{C}^{(N)}$ are the zeros of the $N$th paraorthogonal polynomial (see (1.2))

$$\Phi_N(z, d\mu, \beta) = z\Phi_{N-1}(z, d\mu) - \overline{\beta}\,\Phi_{N-1}^*(z, d\mu) \tag{6.4}$$

Therefore we can consider the complex function

$$B_N(z) = \frac{\beta\, z\, \Phi_{N-1}(z)}{\Phi_{N-1}^*(z)}, \tag{6.5}$$

which has the property that $\Phi_N(e^{i\theta}) = 0$ if and only if $B_N(e^{i\theta}) = 1$.

By writing the polynomials $\Phi_{N-1}$ and $\Phi_{N-1}^*$ as products of their zeros, we can see that the function $B_N$ is a Blaschke product.

Let $\eta_N : [0, 2\pi) \to \mathbb{R}$ be a continuous function such that

$$B_N(e^{i\theta}) = e^{i\,\eta_N(\theta)} \tag{6.6}$$

(we will only be interested in the values of the function $\eta_N$ near a fixed point $e^{i\theta_0} \in \partial\mathbb{D}$). Note that for any fixed $\theta \in \partial\mathbb{D}$, we have that $\eta(\theta)$ is a random variable depending on $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{N-2}, \alpha_{N-1} = \beta) \in \Omega$.

We will now study the properties of the random variable $\eta_N(\theta) = \eta_N(\theta, \alpha_0, \alpha_1, \ldots, \alpha_{N-2}, \beta)$. Thus

**Lemma 6.1.** *For any $\theta_1$ and $\theta_2$, the random variables $\frac{\partial \eta_N}{\partial \theta}(\theta_1)$ and $\eta_N(\theta_2)$ are independent. Also for any fixed value $w \in \mathbb{R}$,*

$$\mathbb{E}\left( \frac{\partial \eta_N}{\partial \theta}(\theta_1) \,\middle|\, \eta_N(\theta_2) = w \right) = N. \tag{6.7}$$

**Proof.** Eq. (6.5) gives

$$\eta_N(\theta) = \gamma + \tau(\theta), \tag{6.8}$$

where $e^{i\gamma} = \beta$ and $e^{i\tau(\theta)} = \frac{e^{i\theta}\Phi_{N-1}(e^{i\theta})}{\Phi^*_{N-1}(e^{i\theta})}$. Since the distribution of each of the random variables $\alpha_0, \alpha_1, \ldots, \alpha_{N-2}$ and $\beta$ is rotationally invariant, for any $\theta \in [0, 2\pi)$, $\gamma$ and $\tau(\theta)$ are random variables uniformly distributed. Also, it is immediate that $\gamma$ and $\tau(\theta)$ are independent. Since $\gamma$ does not depend on $\theta$, for any fixed $\theta_1, \theta_2 \in [0, 2\pi)$, we have that the random variables $\frac{\partial \eta_N}{\partial \theta}(\theta_1)$ and $\eta_N(\theta_2)$ are independent.

We see now that for any Blaschke factor $B_a(z) = \frac{z-a}{1-\bar{a}z}$, we can define a real-valued function $\eta_a$ on $\partial\mathbb{D}$ such that

$$e^{i\,\eta_a(\theta)} = B_a(e^{i\theta}). \tag{6.9}$$

A straightforward computation gives

$$\frac{\partial \eta_a}{\partial \theta}(\theta) = \frac{1 - |a|^2}{|e^{i\theta} - a|^2} > 0. \tag{6.10}$$

Since $B_N$ is a Blaschke product, we now get that for any fixed $\alpha \in \Omega$, $\frac{\partial \eta_N}{\partial \theta}$ has a constant sign (positive). This implies that the function $\eta_N$ is strictly increasing. The function $B_N(z)$ is analytic and has exactly $N$ zeros in $\mathbb{D}$ and therefore we get, using the argument principle, that

$$\int_0^{2\pi} \frac{\partial \eta_N}{\partial \theta}(\theta)\, d\theta = 2\pi N. \tag{6.11}$$

Note that $\frac{\partial \eta_N}{\partial \theta}$ does not depend on $\beta$ (it depends only on $\alpha_0, \alpha_1, \ldots, \alpha_{N-2}$). Also, using the same argument as in Lemma 3.1, we have that for any angles $\theta$ and $\varphi$,

$$\frac{\partial \eta_N}{\partial \theta}(\theta) = \frac{\partial \tilde{\eta}_N}{\partial \theta}(\theta - \varphi), \tag{6.12}$$

where $\tilde{\eta}$ is the function $\eta$ that corresponds to the Verblunsky coefficients

$$\alpha_{k,\varphi} = e^{-i(k+1)\varphi}\alpha_k, \quad k = 0, 1, \ldots, (N-2). \tag{6.13}$$

Since the distribution of $\alpha_0, \alpha_1, \ldots, \alpha_{N-2}$ is rotationally invariant, we get from (6.12) that the function $\theta \to \mathbb{E}\left(\frac{\partial \eta_N}{\partial \theta}(\theta)\right)$ is constant.

Taking expectations and using Fubini's theorem (as we also did in Lemma 3.1), we get, for any angle $\theta_0$,

$$2\pi N = \mathbb{E}\left(\int_0^{2\pi} \frac{\partial \eta_N}{\partial \theta}(\theta)\, d\theta\right) = \int_0^{2\pi} \mathbb{E}\left(\frac{\partial \eta_N}{\partial \theta}(\theta)\right) d\theta = 2\pi\, \mathbb{E}\left(\frac{\partial \eta_N}{\partial \theta}(\theta_0)\right) \tag{6.14}$$

and therefore

$$\mathbb{E}\left(\frac{\partial \eta_N}{\partial \theta}(\theta_0)\right) = N. \tag{6.15}$$

Since for any $\theta_1, \theta_2 \in [0, 2\pi)$, we have that $\frac{\partial \eta_N}{\partial \theta}(\theta_1)$ and $\eta_N(\theta_2)$ are independent, (6.15) implies that for any fixed value $w \in \mathbb{R}$,

$$\mathbb{E}\left(\frac{\partial \eta_N}{\partial \theta}(\theta_1)\,\middle|\, \eta_N(\theta_2) = w\right) = N. \qquad \square \tag{6.16}$$

We will now control the probability of having at least two eigenvalues in $I_N$ conditioned by the event that we already have an eigenvalue at one fixed point $e^{i\theta_1} \in I_N$. This will be shown in the following lemma:

**Lemma 6.2.** *With $\mathcal{C}^{(N)}$, $I_N$, and the events $A(m, \mathcal{C}, I)$ and $M(e^{i\theta})$ defined before, and for any $e^{i\theta_1} \in I_N$, we have*

$$\mathbb{P}\left(A\left(2, \mathcal{C}^{(N)}, I_N\right) \mid M(e^{i\theta_1})\right) \leqslant (b-a). \tag{6.17}$$

**Proof.** Using the fact that the function $\theta \to \mathbb{E}\left(\frac{\partial \eta_N}{\partial \theta}(\theta)\right)$ is constant and the relation (6.16), we get that

$$\mathbb{E}\left(\int_{\theta_0+\frac{2\pi a}{N}}^{\theta_0+\frac{2\pi b}{N}} \frac{\partial \eta_N}{\partial \theta}(\theta_1)\, d\theta_1 \,\middle|\, \eta_N(\theta_2) = w\right) = 2\pi(b-a). \tag{6.18}$$

We see that

$$\Phi_N(e^{i\theta}) = 0 \iff B_N(e^{i\theta}) = 1 \iff \eta_N(\theta) = 0 \pmod{2\pi}. \tag{6.19}$$

Therefore if the event $A(2, \mathcal{C}^{(N)}, I_N)$ takes place (i.e., if the polynomial $\Phi_N$ vanishes at least twice in the interval $I_N$), then the function $\eta_N$ changes by at least $2\pi$ in the interval $I_N$, and therefore we have that

$$\int_{\theta_0+\frac{2\pi a}{N}}^{\theta_0+\frac{2\pi b}{N}} \frac{\partial \eta_N}{\partial \theta}(\theta)\, d\theta \geqslant 2\pi, \tag{6.20}$$

whenever the event $A(2, \mathcal{C}^{(N)}, I_N)$ takes place.

For any $\theta_1 \in I_N$ we have, using the independence of the random variables $\frac{\partial \eta_N}{\partial \theta}(\theta_1)$ and $\eta_N(\theta_2)$ for the first inequality and Chebyshev's inequality for the second inequality,

$$\begin{aligned}
&\mathbb{P}\left(A(2,\mathcal{C}^{(N)},I_N)\,\Big|\, M(e^{i\theta_1})\right)\\
&\leqslant \mathbb{P}\left(\int_{\theta_0+\frac{2\pi a}{N}}^{\theta_0+\frac{2\pi b}{N}} \frac{\partial \eta_N}{\partial \theta}(\theta)\,d\theta \geqslant 2\pi \,\Big|\, M(e^{i\theta_1})\right)\\
&\leqslant \frac{1}{2\pi}\,\mathbb{E}\left(\int_{\theta_0+\frac{2\pi a}{N}}^{\theta_0+\frac{2\pi b}{N}} \frac{\partial \eta_N}{\partial \theta}(\theta)\,d\theta \,\Big|\, M(e^{i\theta_1})\right).
\end{aligned} \tag{6.21}$$

The previous formula shows that we can control the probability of having more than two eigenvalues in the interval $I_N$ conditioned by the event that a fixed $e^{i\theta_1}$ is an eigenvalue. We now obtain, using (6.18) with $w = 2\pi m$, $m \in \mathbb{Z}$,

$$\mathbb{P}\left(A\left(2,\mathcal{C}^{(N)},I_N\right) \mid M\left(e^{i\theta_1}\right)\right) \leqslant (b-a). \qquad \square \tag{6.22}$$

We can now control the probability of having two or more eigenvalues in $I_N$.

**Theorem 6.3.** *With $\mathcal{C}^{(N)}$, $I_N$, and the event $A(m,\mathcal{C},I)$ defined before, we have*

$$\mathbb{P}\left(A\left(2,\mathcal{C}^{(N)},I_N\right)\right) \leqslant \frac{(b-a)^2}{2}. \tag{6.23}$$

**Proof.** For any positive integer $k$, we have

$$\mathbb{P}\left(B(k,\mathcal{C}^{(N)},I_N)\right) = \frac{1}{k}\int_{\theta_0+\frac{2\pi a}{N}}^{\theta_0+\frac{2\pi b}{N}} \mathbb{P}\left(B(k,\mathcal{C}^{(N)},I_N) \mid M(e^{i\theta})\right) N\,d\nu_N(\theta) \tag{6.24}$$

(where the measure $\nu_N$ is the density of eigenvalues).

Note that the factor $\frac{1}{k}$ appears because the selected point $e^{i\theta}$ where we take the conditional probability can be any one of the $k$ points.

We will now use the fact that the distribution of the Verblunsky coefficients is rotationally invariant and therefore for any $N$ we have $d\nu_N = \frac{d\theta}{2\pi}$, where $\frac{d\theta}{2\pi}$ is the normalized Lebesgue measure on the unit circle.

Since for any $k \geqslant 2$ we have $\frac{1}{k} \leqslant \frac{1}{2}$, we get that for any integer $k \geqslant 2$ and for large $N$,

$$\mathbb{P}\left(B(k,\mathcal{C}^{(N)},I_N)\right) \leqslant \frac{N}{2}\int_{\theta_0+\frac{2\pi a}{N}}^{\theta_0+\frac{2\pi b}{N}} \mathbb{P}\left(B(k,\mathcal{C}^{(N)},I_N) \mid M(e^{i\theta})\right) \frac{d\theta}{2\pi} \tag{6.25}$$

and therefore,

$$\mathbb{P}\left(A(2,\mathcal{C}^{(N)},I_N)\right) \leqslant \frac{N}{2}\int_{\theta_0+\frac{2\pi a}{N}}^{\theta_0+\frac{2\pi b}{N}} \mathbb{P}\left(A(2,\mathcal{C}^{(N)},I_N) \mid M(e^{i\theta})\right) \frac{d\theta}{2\pi}. \tag{6.26}$$

Using Lemma 6.2, we get

$$\mathbb{P}(A(2, \mathcal{C}^{(N)}, I_N)) \leqslant \frac{N}{2} \, \frac{(b-a)}{N} \, (b-a) = \frac{(b-a)^2}{2}. \qquad \square \tag{6.27}$$

**Theorem 6.4.** *With $\mathcal{C}^{(N)}, \mathcal{C}_1^{(N)}, \mathcal{C}_2^{(N)}, \ldots, \mathcal{C}_{[\ln n]}^{(N)}, I_N$, and the event $A(m, \mathcal{C}, I)$ defined before, we have, for any $k$, $1 \leqslant k \leqslant [\ln n]$,*

$$\mathbb{P}\left(A(2, \mathcal{C}_k^{(N)}, I_N)\right) = O\left(([\ln n])^{-2}\right) \quad \text{as} \quad n \to \infty. \tag{6.28}$$

**Proof.** We will use the previous theorems for the CMV matrix $\mathcal{C}_k^{(N)}$. Recall that $N = [\ln n] \left[\frac{n}{\ln n}\right]$. Since this matrix has $\left[\frac{n}{\ln n}\right]$ eigenvalues, we can use the proof of Lemma 6.2 to obtain that for any $e^{i\theta} \in I_N$,

$$\mathbb{P}\left(A(2, \mathcal{C}_k^{(N)}, I_N) \,\middle|\, M(e^{i\theta})\right) \leqslant \frac{1}{2\pi} \, \frac{2\pi(b-a)}{N} \left[\frac{n}{\ln n}\right] = \frac{b-a}{[\ln n]}. \tag{6.29}$$

The proof of Theorem 6.3 now gives

$$\mathbb{P}\left(A(2, \mathcal{C}_k^{(N)}, I_N)\right) \leqslant \frac{(b-a)^2}{2\,[\ln n]^2} \tag{6.30}$$

and hence (6.28) follows. $\square$

This theorem shows that as $N \to \infty$, any of the decoupled matrices contributes with at most one eigenvalue in each interval of size $\frac{1}{N}$.

## 7. Proof of the main theorem

We will now use the results of Sections 3–6 to conclude that the statistical distribution of the zeros of the random paraorthogonal polynomials is Poisson.

**Proof of Theorem 1.1.** It is enough to study the statistical distribution of the zeros of polynomials of degree $N = [\ln n] \left[\frac{n}{\ln n}\right]$. These zeros are exactly the eigenvalues of the CMV matrix $\mathcal{C}^{(N)}$, so, by the results in Section 5, the distribution of these zeros can be approximated by the distribution of the direct sum of the eigenvalues of $[\ln n]$ matrices $\mathcal{C}_1^{(N)}, \mathcal{C}_2^{(N)}, \ldots, \mathcal{C}_{[\ln n]}^{(N)}$.

In Section 6 (Theorem 6.4), we showed that the probability that any of the matrices $\mathcal{C}_1^{(N)}, \mathcal{C}_2^{(N)}, \ldots, \mathcal{C}_{[\ln n]}^{(N)}$ contributes with two or more eigenvalues in each interval of size $\frac{1}{N}$ situated near a fixed point $e^{i\theta} \in \partial\mathbb{D}$ is of order $O([\ln n]^{-2})$. Since the matrices $\mathcal{C}_1^{(N)}, \mathcal{C}_2^{(N)}, \ldots,$ $\mathcal{C}_{[\ln n]}^{(N)}$ are identically distributed and independent, we immediately get that the probability that the direct sum of these matrices has two or more eigenvalues in an interval of size $\frac{1}{N}$ situated near $e^{i\theta}$ is $[\ln n]\, O([\ln n]^{-2})$ and therefore converges to 0 as $n \to \infty$.

We can now conclude that as $n \to \infty$, the local distribution of the eigenvalues converges to a Poisson process with intensity measure $n\,\frac{d\theta}{2\pi}$ using a standard technique in probability theory. We first fix an interval $I_N = (e^{i(\theta_0 + \frac{2\pi a}{N})}, e^{i(\theta_0 + \frac{2\pi b}{N})})$ near the point $e^{i\theta_0}$ (as before, we take $N = [\ln n]\left[\frac{n}{\ln n}\right]$). Let us consider $[\ln n]$ random variables $X_1, X_2, \ldots, X_{[\ln n]}$ where $X_k$ = number of the eigenvalues of the matrix $\mathcal{C}_k^{(N)}$ situated in the interval $I_N$ and let $S_n(I_N) = X_1 + X_2 + \cdots + X_{[\ln n]}$. Note that $S_n(I_N)$ = the number of eigenvalues of the matrix $\tilde{\mathcal{C}}^{(N)}$ situated in the interval $I_N$. We want to prove that

$$\lim_{n\to\infty} \mathbb{P}(S_n(I_N) = k) = e^{-(b-a)}\,\frac{(b-a)^k}{k!}. \tag{7.1}$$

Theorem 6.4 shows that we can assume without loss of generality that for any $k$, $1 \leqslant k \leqslant [\ln n]$, we have $X_k \in \{0, 1\}$. Also, because of rotation invariance, we can assume, for large $n$,

$$\mathbb{P}(X_k = 1) = \frac{(b-a)}{[\ln n]}, \tag{7.2}$$

$$\mathbb{P}(X_k = 0) = 1 - \frac{(b-a)}{[\ln n]}. \tag{7.3}$$

The random variable $S_n(I_N)$ can now be viewed as the sum of $[\ln n]$ Bernoulli trials, each with the probability of success $\frac{(b-a)}{[\ln n]}$ and

$$\mathbb{P}(S_n(I_N) = k) = \binom{[\ln n]}{k}\left(\frac{(b-a)}{[\ln n]}\right)^k \left(1 - \frac{(b-a)}{[\ln n]}\right)^{[\ln n]-k} \tag{7.4}$$

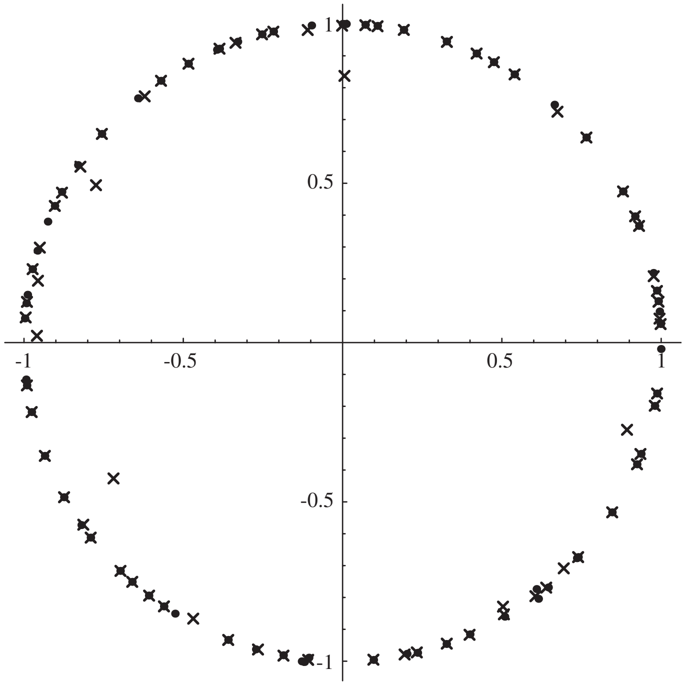
which converges to $e^{-\lambda}\frac{\lambda^k}{k!}$, where $\lambda = [\ln n]\frac{(b-a)}{[\ln n]} = (b-a)$. Therefore we get (7.1).

Since for any disjoint intervals $I_{N,k}$, $1 \leqslant k \leqslant [\ln n]$ situated near $e^{i\theta_0}$, the random variables $S_n(I_{N,k})$ are independent, (7.1) will now give (1.6) and therefore the proof of the main theorem is complete. □

## 8. Remarks

1. We should emphasize the fact that the distribution of our random Verblunsky coefficients is rotationally invariant. This assumption is used in several places and seems vital for our approach. It is not clear how (or whether) the approach presented here can be extended to distributions that are not rotationally invariant.
2. In this paper, we study the statistical distribution of the zeros of paraorthogonal polynomials. It would be interesting to understand the statistical distribution of the zeros of orthogonal polynomials. A generic plot of the zeros of paraorthogonal polynomials

versus the zeros of orthogonal polynomials is




In this Mathematical plot, the points represent the zeros of paraorthogonal polynomials obtained by randomly choosing $\alpha_0, \alpha_1, \ldots, \alpha_{69}$ from the uniform distribution on $D(0, \frac{1}{2})$ and $\alpha_{70}$ from the uniform distribution on $\partial\mathbb{D}$. The crosses represent the zeros of the orthogonal polynomials obtained from the same $\alpha_0, \alpha_1, \ldots, \alpha_{69}$ and an $\alpha_{70}$ randomly chosen from the uniform distribution on $D(0, \frac{1}{2})$.

We observe that, with the exception of a few points (corresponding probably to "bad eigenfunctions"), the zeros of paraorthogonal polynomials and those of orthogonal polynomials are very close. We conjecture that these zeros are pairwise exponentially close with the exception of $O((\ln N)^2)$ of them. We expect that the distribution of the arguments of the zeros of orthogonal polynomials on the unit circle is also Poisson.

3. We would also like to mention the related work of Bourget et al. [5] and Joye [21,22]. In these papers, the authors analyze the spectral properties of a class of five-diagonal random unitary matrices similar to the CMV matrices (with the difference that it contains an extra random parameter). In [22] (a preprint which appeared as this work was being completed), the author considers a subclass of the class of Bourget et al. [5] that does not overlap with the orthogonal polynomials on the unit circle and proves Aizenman–Molchanov bounds similar to the ones we have in Section 3.
4. The results presented in our paper were announced by Simon in [32], where he describes the distribution of the zeros of orthogonal polynomials on the unit circle in two distinct (and, in a certain way, opposite) situations. In the first case, of random Verblunsky coefficients, our paper shows that there is no local correlation between the zeros (Poisson

behavior). The second case consists of Verblunsky coefficients given by the formula $\alpha_n = Cb^n + O((b\Delta)^n)$ (i.e., $\alpha_n/b^n$ converges to a constant $C$ sufficiently fast). In this case it is shown in [32] that the zeros of the orthogonal polynomials are equally spaced on the circle of radius $b$, which is equivalent to saying that the angular distance between nearby zeros is $2\pi/n$ ("clock" behavior).

## Acknowledgments

I am very grateful to B. Simon for his help and encouragement in the investigation of this problem. I am also grateful to M. Aizenman for the personal communication [1].

I thank the anonymous referee for making many helpful comments on the original manuscript, as well as for pointing out Refs. [14,15,34] (from the orthogonal polynomials literature) and [10–12,19,27] (from the signal processing and data analysis literature).

## References

[1] M. Aizenman, personal communication.

[2] M. Aizenman, S. Molchanov, Localization at large disorder and at extreme energies: an elementary derivation, Comm. Math. Phys. 157 (2) (1993) 245–278.

[3] M. Aizenman, J.H. Schenker, R.M. Friedrich, D. Hundertmark, Finite-volume fractional-moment criteria for Anderson localization, Dedicated to Joel L. Lebowitz, Comm. Math. Phys. 224 (1) (2001) 219–253.

[4] P.W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. 109 (1958) 1492–1505.

[5] O. Bourget, J.S. Howland, A. Joye, Comm. Math. Phys. 234 (2) (2003) 191–227.

[6] M.J. Cantero, L. Moral, L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, Linear Algebra Appl. 362 (2003) 29–56.

[7] D.J. Daley, D. Vere-Jones, An Introduction to the Theory of Point Processes, Springer Series in Statistics, Springer, New York, 1988.

[8] R. del Rio, S. Jitomirskaya, Y. Last, B. Simon, Operators with singular continuous spectrum, IV: Hausdorff dimensions, rank one perturbations, and localization, J. Anal. Math. 69 (1996) 153–200.

[9] P.L. Duren, Theory of $H^p$ Spaces, Pure and Applied Mathematics, vol. 38, Academic Press, New York, London, 1970.

[10] J.D. Fournier, Complex zeros of random Szegö polynomials, in: N. Papamichael, S. Ruscheweyh, E.B. Saff (Eds.), Computational Methods and Function Theory CMFT97, World Scientific, Singapore, 1999, pp. 203–223.

[11] J.D. Fournier, G. Mantica, A. Mezincescu, D. Bessis, Universal statistical behavior of the complex zeros of the Wiener transfer functions, Europhys. Lett. 22 (5) (1993) 325–331.

[12] J.D. Fournier, G. Mantica, A. Mezincescu, D. Bessis, Statistical properties of the zeros of the transfer functions in signal processing, in: D. Benest, C. Froesché (Eds.), Chaos and Diffusion in Hamiltonian Systems, Publication Editions Frontières, Orsay, 1996, pp. 257–271.

[13] J. Fröhlich, T. Spencer, Absence of diffusion in the Anderson tight binding model for large disorder or low energy, Comm. Math. Phys. 88 (2) (1983) 151–184.

[14] J.S. Geronimo, Methods of Approximation Theory in Complex Analysis and Mathematical Physics (Leningrad, 1991), Lecture Notes in Mathematics, vol. 1550, Springer, Berlin, 1993, pp. 43–61.

[15] J.S. Geronimo, A. Teplyaev, A difference equation arising from the trigonometric moment problem having random reflection coefficients—an operator-theoretic approach, J. Funct. Anal. 123 (1) (1994) 12–45.

[16] I.Ya. Goldsheid, S.A. Molchanov, L.A. Pastur, A random homogeneous Schrödinger operator has a pure point spectrum, Funkcional. Anal. i Prilovzen. 11(1) (1977) 1–10, 96 (in Russian). (Engl. Transl.: Funct. Anal. Appl. 11(1) (1977) 1–8.)

[17] L. Golinskii, Quadrature formula and zeros of para-orthogonal polynomials on the unit circle, Acta Math. Hungar. 96 (3) (2002) 169–186.
[18] L. Golinskii, P. Nevai, Szegö difference equations, transfer matrices and orthogonal polynomials on the unit circle, Comm. Math. Phys. 223 (2) (2001) 223–259.
[19] W.B. Jones, O. NjÅstad, E.B. Saff, Szegö polynomials associated with Wiener–Levinson filters, J. Comput. Appl. Math. 32 (3) (1990) 387–406.
[20] W.B. Jones, O. Njastad, W.J. Thron, Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle, Bull. London Math. Soc. 21 (2) (1989) 113–152.
[21] A. Joye, Density of states and Thouless formula for random unitary band matrices, Ann. H. Poincaré 5 (2) (2004) 347–379.
[22] A. Joye, Fractional moment estimates for random unitary operators, preprint.
[23] A. Khodakovsky, Inverse spectral problem with partial information on the potential, Ph.D. Thesis, Caltech, 1999.
[24] S. Khrushchev, Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in $L^2(\mathbb{T})$, J. Approx. Theory 108 (2001) 161–248.
[25] N. Minami, Local fluctuation of the spectrum of a multidimensional Anderson tight binding model, Comm. Math. Phys. 177 (3) (1996) 709–725.
[26] S.A. Molchanov, The local structure of the spectrum of the one-dimensional Schrödinger operator, Comm. Math. Phys. 78 (3) (1980/81) 429–446.
[27] L. Pakula, Asymptotic zero distribution of orthogonal polynomials in sinusoidal frequency estimation, IEEE Trans. Inform. Theory 33 (4) (1987) 569–576.
[28] B. Simon, Spectral Analysis of Rank One Perturbations and Applications, Mathematical Quantum Theory, II: Schrödinger Operators (Vancouver, BC, 1993), CRM Proceedings of the Lecture Notes, vol. 8, American Mathematical Society, Providence, RI, 1995, pp. 109–149.
[29] B. Simon, Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory, American Mathematical Society, Providence, RI, 2005.
[30] B. Simon, Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory, American Mathematical Society, Providence, RI, 2005.
[31] B. Simon, Aizenman's theorem for orthogonal polynomials on the unit circle, Const. Approx., to appear.
[32] B. Simon, Fine structure of the zeros of orthogonal polynomials, I: a tale of two pictures, preprint.
[33] B. Simon, T. Wolff, Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians, Comm. Pure Appl. Math. 39 (1) (1986) 75–90.
[34] A.V. Teplyaev, Orthogonal polynomials with random recurrence coefficients, Trans. Sci. Sem. LOMI 17 (1989) 157–162.
[35] A.V. Teplyaev, The pure point spectrum of random orthogonal polynomials on the circle, Dokl. Akad. Nauk SSSR 320 (1) (1991) 49–53 trans. Sov. Math. Dokl. 44 (2) (1992) 407–411.