DIMENSION RESULTS FOR THE SPECTRAL MEASURE OF THE CIRCULAR β ENSEMBLES

BY TOM ALBERTS[1,a] AND RAOUL NORMAND[2,b]

[1]*Department of Mathematics, University of Utah,* [a]*alberts@math.utah.edu*

[2]*Courant Institute of Mathematical Sciences, New York University,* [b]*rjn5@nyu.edu*

We study the dimension properties of the spectral measure of the circular β-ensembles. For $\beta \geq 2$ it was previously shown by Simon that the spectral measure is almost surely singular continuous with respect to Lebesgue measure on $\partial \mathbb{D}$ and the dimension of its support is $1 - 2/\beta$. We reprove this result with a combination of probabilistic techniques and the so-called Jitomirskaya–Last inequalities. Our method is simpler in nature and mostly self-contained, with an emphasis on the probabilistic aspects rather than the analytic. We also extend the method to prove a large deviations principle for norms involved in the Jitomirskaya–Last analysis.

1. Introduction.

1.1. *Circular β ensemble.* The circular β-ensemble $C\beta E_n$ is the point process of n random points on the unit circle that are distributed according to the law

$$(1.1) \qquad \frac{1}{Z_{n,\beta}} \prod_{1 \leq i < j \leq n} |e^{i\theta_i} - e^{i\theta_j}|^\beta \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi}.$$

The partition function $Z_{n,\beta}$ normalizes the density to be a probability measure and is known explicitly [23, 44]. The circular β-ensemble was introduced by Dyson [17] who also observed that its density is the Gibbs measure for n identical charged particles confined to the unit circle and interacting via the two-dimensional Coulombic repulsion. For this reason the point process is also sometimes called a log-gas. It is also known that (1.1) is the stationary density for the Dyson Brownian motion on $\partial \mathbb{D}$, with the strength of the interaction determined by β.

That the points live on the unit circle suggests that they should be the eigenvalues of a random unitary matrix, and this is well known in the $\beta = 2$ case (the circular unitary ensemble). For general β, the corresponding random matrices were constructed in [28] using the so-called *CMV representation* [10, 29, 39]. CMV is a very general yet efficient matrix representation of a unitary operator on a Hilbert space, analogous to the representation of a self-adjoint operator by a Jacobi matrix. The CMV representation is an infinite matrix that can be seen as an operator on $\ell^2(\mathbb{N})$, parameterized by an infinite sequence $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots)$ taking values in $\overline{\mathbb{D}}$, the closure of the unit disk. These coefficients are now commonly referred to as the *Verblunsky coefficients*, and the correponding matrix denoted $\mathcal{C}(\boldsymbol{\alpha})$. To construct a random unitary matrix whose eigenvalue distribution is the circular β-ensemble for general $\beta > 0$, it is therefore enough to find an appropriate sequence of *random* Verblunsky coefficients, and construct the corresponding CMV matrices. This is the approach successfully used in [28], though we will use the slightly different choice of Verblunskys made in [30]. They prove that the appropriate choice is to take the α_j independent with law given by

$$(1.2) \qquad \alpha_j \sim e^{2\pi i U} \sqrt{\text{Beta}\left(1, \frac{\beta}{2}(j + 1)\right)},$$

where U is a Uniform random variable on $(0, 1)$ that is independent of the Beta variable, and all are independent across j. Another way to put it is that the pdf of α_j with respect to the Lebesgue measure on $\mathbb{D}$ is

$$f_j(z) = \frac{\beta(j+1)}{2\pi}(1 - |z|^2)^{\beta(j+1)/2-1}.$$

Note that these variables are independent under all rotations and hence have mean zero, while the second moment of their radial part decays like

$$(1.3) \qquad \mathbb{E}[|\alpha_k|^2] = \frac{2}{\beta(k+1)+2} \sim \frac{2}{\beta k}.$$

It is shown in [28, 30] that if U_0 is another independent uniform$(0, 1)$ random variable then the CMV matrix

$$(1.4) \qquad \mathcal{C}_n := \mathcal{C}(\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, e^{2\pi i U_0}, 0, 0, 0, \ldots)$$

has (1.1) as its eigenvalue density. We describe the CMV matrices more precisely in Section 2.1, but for now one can think of them as infinite matrices. It will follow that $\mathcal{C}_n$ has zeros on the rows and columns of order greater than n and its $n \times n$ minor is a matrix with eigenvalue distribution given by (1.2). Moreover, these operators are "nested" within each other as n increases. The operator $\mathcal{C}_{n+1}$ is obtained from $\mathcal{C}_n$ by simply shifting the $e^{2\pi i U_0}$ entry up by one and inserting α_{n-1} into the now empty slot. In this sense $\mathcal{C}_n$ is a minor of $\mathcal{C}_{n+1}$, and in fact all $\mathcal{C}_n$ are minors of the infinite matrix $\mathcal{C}(\boldsymbol{\alpha})$. This feature is present in the CMV matrices used in [30] but not those introduced in [28], which is the reason behind our choice of these Verblunskys.

As $\boldsymbol{\alpha}$ is a random family of Verblunsky coefficients in $\mathbb{D}$, the operator $\mathcal{C}(\boldsymbol{\alpha})$ is a random operator on $\ell^2(\mathbb{N})$, and it can be shown that it has $(1, 0, 0, \ldots)$ as a cyclic vector. It therefore has a spectral measure $\mu_{\boldsymbol{\alpha}}$ on $\partial\mathbb{D}$, which encodes information about both the eigenvalues and eigenvectors of the operator. As the sequence $\boldsymbol{\alpha}$ is random, it therefore gives rise to a random measure in the usual sense of the term; see [26]. This is what we shall call the *spectral measure of the circular β-ensemble*. Note that the choice of Verblunksy coefficients from [28] would have a different spectral measure, despite having the same eigenvalue distribution. This is because the spectral weights attached to the eigenvalues are different.

The spectral measure is also intimately connected to the *Liouville quantum gravity measure* on $\partial\mathbb{D}$, as was recently proved in [13]. Liouville quantum gravity has been of intense interest in probability theory and statistical mechanics in recent years for its relation to Schramm–Loewner evolution, conformal field theory, and other parts of the broader field of random conformal geometry (see [2, 19, 33, 35, 42] among others). Perhaps its simplest description is via Gaussian multiplicative chaos, where one begins with a mean zero, log-correlated Gaussian field X on $\partial\mathbb{D}$ with $\mathrm{Cov}(X(e^{i\theta}), X(e^{i\phi})) = -2\log|e^{i\theta} - e^{i\phi}|$ and defines the GMC measure as the following limit, in the topology of weak convergence of measures:

$$e^{\frac{\gamma}{2}X(e^{i\theta})}\,d\theta := \lim_{\epsilon \to 0} e^{\frac{\gamma}{2}X_\epsilon(e^{i\theta}) - \frac{\gamma^2}{8}E[X_\epsilon(e^{i\theta})^2]}\,d\theta,$$

where X_ϵ is an appropriately smoothed out version of X that converges to X in an appropriate sense as $\epsilon \to 0$. The limiting measure is shown to exist and be nonzero for all $\gamma \in [0, 2)$, although the notation for it is somewhat abusive since the limiting measure is almost surely singular with respect to Lebesgue so long as $\gamma > 0$. Most notably, the limiting measure is not a probability measure on $\partial\mathbb{D}$ and its total mass is a random variable (see [13, 19, 35] for recent calculations of its density), in contrast with the spectral measure of the circular β-ensemble which is a random probability measure for all $\beta > 0$. In particular, for the GMC the measure and the total mass go to zero as $\gamma \uparrow 2$, while the spectral measure of circular β is

singular continuous for $\beta \geq 2$ but undergoes a phase transition to an atomic measure as $\beta \downarrow 2$. To observe a similar phase transition for the GMC requires subtle renormalizations as $\gamma \uparrow 2$ (see [1]) which are not needed for circular β, thanks to the theory of orthogonal polynomials on the unit circle.

In the subcritical phase $\gamma < 2$ and the critical phase $\gamma = 2$ ($\beta \geq 2$) Chhaibi and Najnudel [13] recently proved an exact equality between GMC and the spectral measure of the circular-β ensemble. More precisely, they determine an explicit real-valued function of the Verblunsky coefficients whose product with the spectral measure is *exactly* the GMC measure on $\partial \mathbb{D}$, with the relationship $\gamma^2 = 8/\beta$. Note that [13] is actually that $\gamma^2 = 2/\beta$, but this is because they use $e^{\gamma X}$ as their normalization rather than our $e^{(\gamma/2)X}$ (with the same 2 in the covariance kernel of the X). One of our reasons for this is that it maintains the phase transition at $\gamma = 2$ as in [15], but we also note that our normalization is also what is used in [35]. Also note that the result of [13] is somewhat similar to the results of Webb [42] and Nikula and al [34], who prove that (normalized) powers of the characteristic polynomial of random unitary matrices converge to GMC on $\partial \mathbb{D}$. The relation $\gamma^2 = 8/\beta$ agrees with the phase transition occurring at both $\gamma = 2$ and $\beta = 2$. Furthermore, the relation $\gamma = \sqrt{\kappa}$ between GMC and SLE processes [15] leads to $\beta = 8/\kappa$, which was earlier conjectured by Cardy [11] in his relation of Dyson's Brownian motion to multiple SLE. Following the analysis in [3, 15, 24] the GMC measure is supported on a set of Hausdorff dimension $1 - \gamma^2/4$, which by the identification above is equal to $1 - 2/\beta$. One of our main results is a new proof of this dimension formula using the tools of orthogonal polynomials on the unit circle and ideas from probability theory, which we describe next.

1.2. *Results.* In this paper we focus on the almost sure dimension properties of the spectral measure of the circular β-ensemble. We reprove a result of Simon [40], Theorem 12.7.7 (see also [7]) on the almost sure Hausdorff dimension of the support of the random measure by combining the technique of *size-biasing* with the *Jitomirskaya–Last inequalities* for computing local dimensions via orthogonal polynomials. Recently, Simon's result was invoked in [13], Corollary 2.4, to give an alternative proof for the dimension of the support of the Gaussian multiplicative chaos via the direct link they establish to the spectral measure of the circular β-ensemble. As we explain shortly, in the background of Simon's result is an appeal to the abstract but *deterministic* theory of Kotani–Ushiroya [40], Theorem 10.5.34, which is quite involved. Our methods bypass the Kotani–Ushiroya theory and instead use probabilistic ideas to prove the same result as [40], Theorem 12.7.7, in the case of certain random Verblunsky coefficients and with minor modifications to the assumptions. We also prove a large deviation principle that may be relevant to studying the fine dimensional properties of the measure. Additional questions are discussed in [41]. Our results do not rely on the exact form (1.2) for the law of the Verblunsky coefficients. We shall merely use the independence, the rotational symmetry of their laws, and the asymptotics (1.3). To be precise, let us denote by Q the law of the family $(\alpha_0, \alpha_1, \dots)$ on $\mathbb{D}^\infty$, endowed with the product σ-algebra. This gives rise to a random measure $\mu_{\boldsymbol{\alpha}}$ on $\partial \mathbb{D}$. We make the following assumptions, which shall become relevant at different (but clearly stated) parts of the text.

ASSUMPTION 1.1. The Verblunsky coefficients $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots)$ take values in $\mathbb{D}$, are independent, and each have a radially symmetric distribution. In other words, Q is a product measure on $\mathbb{D}^\infty$, and each marginal distribution is invariant under arbitrary rotations of $\mathbb{D}$.

ASSUMPTION 1.2. The second moments of the Verblunsky coefficients decay as

$$(1.5) \qquad\qquad \mathbb{E}\big[|\alpha_k|^2\big] \sim \frac{2}{\beta k}.$$

These are the main assumptions that we will need. The factor β controls the dimension of the measure, as Theorem 1.4 below shows. Note that these two assumptions are clearly satisfied by the Verblunsky coefficients of the circular β-ensemble, which in fact satisfy $\mathbb{E}[|\alpha_k|^m] \sim C_m k^{-m/2}$ as $k \to \infty$ for all positive integers m. In general we do not require such strong asymptotics for the higher moments, but our two main results will also require some additional assumptions on higher moments that rule out the α_k having too much mass near the unit circle. For instance, distributions for $|\alpha_k|$ of the type

$$\left(1 - \frac{1}{k^3}\right)\delta_{k^{-1/2}} + \frac{1}{k^3}\delta_{1-1/k}$$

would not be allowed.

1.2.1. *Exact dimension of the measure.* For our first result recall the following definition of the exact Hausdorff dimension of a measure μ.

DEFINITION 1.3. A measure μ on a measure-metric space has *exact Hausdorff dimension* κ if it is supported on a set of Hausdorff dimension κ, and gives zero mass to all sets of dimension strictly smaller than κ.

THEOREM 1.4. *Assume that the Verblunsky coefficients α_k satisfy Assumptions 1.1 and 1.2 for some $\beta \geq 2$, and that the third moments also satisfy*

$$(1.6) \qquad\qquad E[|\alpha_k|^3] = o(k^{-1}).$$

Then with Q probability one, the exact Hausdorff dimension of μ_α is $1 - 2/\beta$.

A similar result is proved by Simon in [40], Theorem 12.7.7, albeit with slightly different assumptions (see also [6] for similar results for random tridiagonal matrices). Simon also assumes the independence of the Verblunskies and the same decay of the second moment, but instead of rotational invariance, he only assumes the weaker condition that $\mathbb{E}[\alpha_k] = \mathbb{E}[\alpha_k^2] = 0$ for each k. However, he also makes the extra assumptions that

$$(1.7) \qquad\qquad \sup_{n,\omega}|\alpha_n| < 1, \qquad \sup_{n,\omega}\sqrt{n}|\alpha_n| < \infty,$$

which we do not require and is not satisfied for the β-ensemble. Simon also considers the so-called Aleksandrov measures $\mu_{\lambda\alpha}$, where $\lambda \in \partial\mathbb{D}$ and $\lambda\alpha$ means the rotated coefficients $(\lambda\alpha_0, \lambda\alpha_1, \dots)$, and his result is the following.

THEOREM 1.5 ([40], Theorem 12.7.7). *With $dQ(\alpha) d\lambda/2\pi$ probability one the measure $\mu_{\lambda\alpha}$ has exact Hausdorff dimension $1 - 2/\beta$.*

In one sense this is a more general result as no rotational independence is required. On the other hand, when the α_k do have a rotationally symmetric distribution then all the $\mu_{\lambda\alpha}$ have the same distribution, and we therefore recover our Theorem 1.4. However, this comes at the price of Assumptions (1.7) above, which are not satisfied in many situations, most notably for the spectral measure of the CβE. A truncation argument similar to what we apply in Section 4.2 would overcome this difficulty, but as we explain next the main strength of our proof is that it avoids the underlying *Kotani theory* on which Simon's proof is based. This has the advantage of making the proof more self-contained, and highlights the usefulness of the underlying probabilistic ideas.

Both our proofs and Simon's are based on an analysis of the *local dimension* of the measure $\mu_{\boldsymbol{\alpha}}$ at different points. Loosely speaking, the local dimension of a measure ν on $\partial\mathbb{D}$ at a point $e^{i\theta}$ is the exponent $s_0 = s_0(\nu, \theta) \in [0, 1]$ such that

$$\nu(\theta - \epsilon, \theta + \epsilon) \approx \epsilon^{s_0}$$

as $\epsilon \to 0$ (see Section 2.8 for the precise definition). The dimension theory of Rogers and Taylor [37, 38] implies that if the local dimension $\theta \mapsto s_0(\nu, \theta)$ is almost surely constant under the measure ν, then ν has exact Hausdorff dimension equal to this constant. Computing $s_0(\nu, \theta)$ at a particular point usually requires working with some approximating sequence ν_δ of smoothed out versions of the measure, which is often difficult because it requires precise estimates on the relationship between the scale of the approximation δ and the width of the interval ϵ in the quantity $\nu_\delta(\theta - \epsilon, \theta + \epsilon)$. For probability measures on the real line, a particularly nice method of approximation using the associated orthogonal polynomials was developed by Last [32] and Jitomirskaya–Last [25]. Later, Simon [40] translated their results to probability measure on $\partial\mathbb{D}$ and OPUC, which forms the basis of his proof of Theorem 1.5. In fact, this is where we learned of the Jitomirskaya–Last technique. In short, it says that if $\varphi_n : \partial\mathbb{D} \to \mathbb{C}$, $n \geq 0$, are the orthonormal polynomials in the Hilbert space $L^2(\partial\mathbb{D}, d\nu)$ and $\psi_n : \partial\mathbb{D} \to \mathbb{C}$, $n \geq 0$, are the associated *second kind polynomials* (see Section 2.5 for the definition) then

$$s > s_0(\nu, \theta) \quad \Longleftrightarrow \quad \liminf_{n \to \infty} \frac{\|\varphi.(e^{i\theta})\|_n^{2-s}}{\|\psi.(e^{i\theta})\|_n^s} = \infty,$$

where $\|a.\|_n^2$ is the squared norm on sequences $a. = \{a_n\}_{n=0}^{\infty}$ in $\mathbb{C}$ defined by

$$\|a.\|_n^2 = \sum_{k=0}^{n} |a_k|^2.$$

When applied to the orthogonal polynomials we refer to these as the Jitomirskaya–Last norms. Our analysis of the norms is somewhat similar to Simon's but the main difference in our proofs comes in the way we choose θ. Simon appeals to a powerful theorem of Kotani–Ushiroya [40], Theorem 10.5.34, that relates the growth of norms to the existence of points where the measure has a particular local dimension. The advantage of the Kotani–Ushiroya result is that it is purely deterministic, but the underlying theory is difficult and much broader than what is needed for random Verblunsky coefficients. Our approach is briefer and entirely self-contained. It relies only on classical probabilistic ideas: martingale arguments, laws of large numbers, and coupling techniques, and gives a full description of which points θ are the appropriate ones to look at. We analyze the Jitomirskaya–Last norms by considering them as random variables under the joint measure $dQ(\boldsymbol{\alpha}) \, d\mu_{\boldsymbol{\alpha}}(\theta)$, which is a skew product on $\mathbb{D}^{\infty} \times \partial\mathbb{D}$ with the Verblunskies chosen first and then the point θ sampled according to the measure determined by the Verblunskies. We give an alternative description of this measure using the *Bernstein–Szegö approximation*, which says that for each fixed n the quantity $|\varphi_n(e^{i\theta})|^{-2} \, d\theta/2\pi$ is a probability measure on $\partial\mathbb{D}$, and that

$$|\varphi_n(e^{i\theta})|^{-2} \frac{d\theta}{2\pi} \xrightarrow{n \to \infty} d\mu_{\boldsymbol{\alpha}}(\theta)$$

in the sense of weak convergence of measures. The Bernstein–Szegö approximation allows us to rewrite the joint measure $dQ(\boldsymbol{\alpha})d\mu_{\boldsymbol{\alpha}}(\theta)$ in terms of the marginal measure of the point θ and the conditional measure of the Verblunsky coefficients, which turns out to be a much simpler description of the random pair. The conditional measure turns out to be a size-biasing of the original Verblunsky coefficients with nice properties. In particular there is a natural coupling between the original Verblunskies and the size-biased ones that in turn allows us to easily

compute all moments of the size-biased Verblunskies. A similar size-biasing was previously used in [5], although in a slightly different form and for different purposes. In our work the size-biased Verblunskies are used to compute a "strong law of large numbers" for the growth of the Jitormirskaya–Last norms at typical points of the measure μ_{α}, which is then combined with the Rogers–Taylor theory to produce the dimension results.

We point out that for probabilists, the combination of Rogers–Taylor and Jitomirskaya–Last dimension theories is potentially useful in other applications. A now relatively standard method of computing Hausdorff dimensions of random sets is via the so-called *second moment method*. Roughly speaking this requires estimates on the expected value and variance of the number of balls in an ϵ-net of the ambient space that are required to cover the random fractal. The required estimates are on the power law blowup of these quantities as $\epsilon \to 0$. The expected number is typically easy to deal with since it only requires the probability that a fixed point is within a distance ϵ of the random set, but the variance is more complicated since this requires estimates on probabilities of two fixed points both being within distance ϵ of the random set. This type of two point estimate requires getting a handle on often complicated correlations and can quickly become messy. In contrast, the Jitomirskaya–Last dimension theory requires two one-point estimates rather than one two-point estimate, and while these two one-point quantities are related the correlation appears to be simpler to deal with. The necessary estimate is handled in Section 4.5. The downside of course is that Jitomirskaya–Last requires there to be a measure on the random set and some information about the associated orthogonal polynomials, which may not always be available.

1.2.2. *Large deviations for the Jitomirskaya–Last norm.*

Our second main result is a large deviations principle (LDP) for the growth of the Jitomirskaya–Last norms. Given that the dimension result of Theorem 1.4 is based on a strong law for the growth of these norms, an LDP is a very natural extension that should have applications in analyzing the subsets of $\partial \mathbb{D}$ where the measure has atypical local dimension.

THEOREM 1.6. *Assume the Verblunsky coefficients α_k satisfy Assumptions* 1.1 *and* 1.2 *and that, for some $\epsilon > 0$ and all $\kappa > 0$,*

$$(1.8) \qquad \mathbb{E}\big[|\alpha_k|^3\big] = O\big(k^{-1-\epsilon}\big), \qquad \limsup_{k \to \infty} \mathbb{E}\big[(1 - |\alpha_k|)^{-\kappa}\big] < \infty.$$

Then under the measure $d\mathbf{P}(\boldsymbol{\alpha}, \theta) = dQ(\boldsymbol{\alpha}) \, d\theta / 2\pi$, the sequence of random variables

$$\frac{\log \|\varphi_\cdot(e^{i\theta})\|_n^2}{\log n}$$

satisfies a large deviations principle with speed $\log n$ and rate function

$$I(x) = \begin{cases} \dfrac{\beta}{8}\Big(x - 1 - \dfrac{2}{\beta}\Big)^2, & x \geq 0, \\ \infty, & x < 0. \end{cases}$$

Moreover, under the measure $d\mathbb{P}(\boldsymbol{\alpha}, \theta) = dQ(\boldsymbol{\alpha}) \, d\mu_{\boldsymbol{\alpha}}(\theta)$, the same sequence satisfies the LDP with speed $\log n$ and rate function

$$J(x) = \begin{cases} \dfrac{\beta}{8}\Big(x - 1 + \dfrac{2}{\beta}\Big)^2, & x \geq 0, \\ \infty, & x < 0. \end{cases}$$

In particular, these moment assumptions are satisfied by the Verblunsky coefficients of the circular β-ensemble. Indeed, by (1.2), $|\alpha_k|^2$ has a Beta$(1, \beta(k+1)/2)$ distribution, and hence $\mathbb{E}[|\alpha_k|^3] = O(k^{-3/2})$. Moreover, this also implies that $1 - |\alpha_k|^2$ has a Beta$(\beta(k+1)/2, 1)$ distribution, and hence

$$\mathbb{E}[(1 - |\alpha_k|)^{-\kappa}] \leq C\mathbb{E}[(1 - |\alpha_k|^2)^{-\kappa}] = C' \int_0^1 x^{-\kappa} x^{\beta(k+1)/2-1}\, dx = \frac{C''}{\beta(k+1)/2 - \kappa},$$

for some constants C', C'', provided that $\kappa < \beta(k+1)/2$. Note however that the assumptions of Theorem 1.6 means that the quadratic nature of the rate function is not particular to the circular-β ensemble and is therefore universal for a wide class of Verblunsky coefficients, even though the spectral measure itself is in bijection with the Verblunskies. The quadratic rate function is of course a manifestation of the underlying Gaussian field [13] behind the orthogonal polynomials.

The jump in the rate functions at $x = 0$ is because the range $x < 0$ would correspond to a shrinking of $\|\varphi.(1)\|_n^2$ as $n \to \infty$, which is clearly impossible since the norm is the sum of positive terms. That the rate function J has a zero at $x = 1 - 2/\beta$ suggests that $\log \|\varphi.(1)\|_n^2 \sim (1 - 2/\beta) \log n$ almost surely under $\mathbb{P}$, which is a key part of the proof of Theorem 1.4. The large deviations principle can be seen as a refinement of this result, and its proof also contains more detailed information. In particular, in the proof of Theorem 1.4, we only need to consider the almost sure asymptotics of $\log |\varphi_n(e^{i\theta})|^2 / \log n$ which are in turn enough to control the asymptotics of $\log \|\varphi.(e^{i\theta})\|_n^2 / \log n$. But in the large deviations regime, we need to control the behavior of the entire process $k \mapsto \log |\varphi_k(e^{i\theta})|^2$ in order to control the behavior of $\log \|\varphi.(e^{i\theta})\|_n^2$. We do so by proving a process level large deviations principle for $k \mapsto \log |\varphi_k(e^{i\theta})|^{-2}$, suitably rescaled and under an appropriate change of the time scale, which takes up the bulk of our proof of Theorem 1.6 in Section 5. The process level LDP that we prove is an extension of Mogulskii's theorem for sums of i.i.d. random variables, although we need to relax the assumption of identical distribution in order to deal with the decaying Verblunsky coefficients. The process level LDP (Theorem 5.1) is useful in its own right since it explains the behavior of the Verblunsky coefficients when $\log |\varphi_n(e^{i\theta})|^{-2}$ takes on an atypical value, that is, whenever the measure has some unusual behavior around θ. Once the process level LDP is proven, converting it into the LDP for the norm is a relatively straightforward application of the contraction principle.

We point out that large deviations in the context of OPUC have become quite popular in recent years [8, 9, 20–22], although more towards sum rules rather than dimension theory. In particular, [20] uses large deviations techniques to give an alternative proof of the Szegö–Verblunsky sum rule for OPUC, with extensions of this idea in the remaining papers. We will not state the sum rule here since it is concerned with the absolutely continuous part of the spectral measure, of which there is none for the types of Verblunsky coefficients that we are considering.

1.3. *Outline.* The outline of this paper is as follows. Section 2 is background on the OPUC theory and dimension theory that we will use throughout. In Section 3, we explain how the Bernstein–Szegö approximation of measures allows us to study the typical points of the spectral measure via a size-biasing of their Verblunsky coefficients. We also explain the coupling procedure that allows us to transform the original Verblunsky coefficients into their size-biased versions, and how this allows us to compute moments of the size-biased quantities. In Section 4 we use this coupling to prove Theorem 1.4 on the exact dimension of the spectral measure of the circular-β ensemble and other measures coming from rotationally invariant Verblunsky coefficients. Finally, in Section 5, we prove the large deviations principle for the Jitomirskaya–Last norm, using an extension of results for process level large deviations.

2. Background for OPUC theory. In this section we review the deterministic theory behind the theory of orthogonal polynomials on the unit circle (OPUC), which will be our main technique to study the measure $\mu_{\boldsymbol{\alpha}}$. The exposition and notation of this section very closely follows that in [39, 40], where the interested reader can find a much more detailed treatment. We emphasize that almost everything in this section is deterministic in nature; the random case is considered in the following section.

2.1. *CMV matrices.* Originally discovered in [10], the CMV matrices are canonical representatives of unitary operators acting on a separable Hilbert space, much like tri-diagonal matrices are canonical representatives for self-adjoint operators. CMV matrices turn out to be five-diagonal but it is known [39] that they are the sparsest possible matrix representations of a unitary operator, in the sense that no fewer number of diagonals can represent all unitary operators.

The CMV matrices are defined as operators on the complex sequence space $\ell^2(\mathbb{N})$, where $\mathbb{N} = \{0, 1, 2, \dots\}$, with the inner product

$$\langle (a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty} \rangle = \sum_{n=0}^{\infty} \overline{a_n} b_n.$$

Note that this inner product is antilinear in the first factor, which is the convention we follow throughout. The input to forming a CMV matrix is an infinite sequence of complex numbers $\{\alpha_n\}_{n=0}^{\infty}$ taking values in the closure $\overline{\mathbb{D}}$ of the unit disk $\mathbb{D}$. For each such sequence we make the following definition.

DEFINITION 2.1. Let $\boldsymbol{\alpha} = \{\alpha_j, j \geq 0\}$ be a sequence taking values in $\overline{\mathbb{D}}$. Let

$$(2.1) \qquad \rho_j = \sqrt{1 - |\alpha_j|^2}$$

and define the 2×2 matrices Θ_j by

$$(2.2) \qquad \Theta_j = \begin{pmatrix} \overline{\alpha}_j & \rho_j \\ \rho_j & -\alpha_j \end{pmatrix}.$$

Let $\mathcal{M}$ and $\mathcal{L}$ be the infinite matrices

$$\mathcal{M} = \begin{pmatrix} 1 & & & \\ & \Theta_1 & & \\ & & \Theta_3 & \\ & & & \ddots \end{pmatrix}, \qquad \mathcal{L} = \begin{pmatrix} \Theta_0 & & & \\ & \Theta_2 & & \\ & & \Theta_4 & \\ & & & \ddots \end{pmatrix},$$

where the 1 in $\mathcal{M}$ is just the number 1, not an identity matrix, and the blank entries are all zero. The Θ_j matrices are always aligned so that the α_j terms lie on the diagonal. Then the *CMV matrix* associated to the sequence $\boldsymbol{\alpha}$ is the infinite matrix (or operator) on $\ell^2(\mathbb{N})$ defined by

$$\mathcal{C}(\boldsymbol{\alpha}) = \mathcal{C}(\alpha_0, \alpha_1, \alpha_2, \dots) := \mathcal{L}\mathcal{M}.$$

It is straightforward to see that the CMV matrices are unitary and it is not much more difficult to compute that they are at most five-diagonal. The complex numbers $(\alpha_n, n \geq 0)$, historically have more than one name in the literature but are now commonly referred to as the *Verblunsky coefficients*. We also use this terminology.

2.2. *Spectral measure.* A key tool in the study of unitary operators is their *spectral measure*, based on the spectral theorem [39] which we now briefly recall. Assume that $U : \mathcal{H} \to \mathcal{H}$ is a unitary operator on a separable Hilbert space $\mathcal{H}$, and further assume that there is a $\zeta \in \mathcal{H}$ such that the finite linear span of $\{U^i\zeta : i \in \mathbb{Z}\}$ is dense in $\mathcal{H}$. Such a ζ is called a *cyclic vector* for U. The spectral theorem asserts that there is a probability measure μ on $\partial\mathbb{D}$ and a linear mapping $V : \mathcal{H} \to L^2(\partial\mathbb{D}, \mathrm{d}\mu)$ such that:

- $V(Ux) = zV(x)$, that is, U on the image side is just multiplication by the *function z*;
- V preserves the inner product, that is, $\langle Vx, Vy \rangle_{L^2(\partial\mathbb{D}, \mathrm{d}\mu)} = \langle x, y \rangle_{\mathcal{H}}$, and hence V is an isometry;
- $V(\zeta) = 1$, the constant function on $\partial\mathbb{D}$.

This μ is called the *spectral measure for the pair* (U, ζ). Since the mapping V preserves the inner product we see that applying U in $\mathcal{H}$ is equivalent to multiplying by z in $L^2(\partial\mathbb{D}, \mathrm{d}\mu)$. Hence studying U is entirely reduced to studying properties of the Hilbert space $L^2(\partial\mathbb{D}, \mathrm{d}\mu)$, and in that sense the spectral measure μ encodes all the information about U.

For the CMV matrices, if the Verblunsky coefficients α_n all satisfy $|\alpha_n| < 1$ then it can be shown (see the next section) that $\delta_0 = (1, 0, 0, \ldots)$ is a cyclic vector for $\mathcal{C}(\boldsymbol{\alpha})$ as an operator on $\ell^2(\mathbb{N})$. To emphasize the dependence on the Verblunsky coefficients we will write $\mu_{\boldsymbol{\alpha}}$ for the spectral measure of the pair $(\mathcal{C}(\boldsymbol{\alpha}), \delta_0)$. The mapping $\boldsymbol{\alpha} \mapsto \mu_{\boldsymbol{\alpha}}$ is measurable from $\mathbb{D}^\infty$ to the space of probability measures on $\partial\mathbb{D}$; which follows from the Bernstein–Szegő approximation (Proposition 2.2).

2.3. *Orthogonal polynomials.* One of the main tools used to study the spectral measure is the associated *orthogonal polynomials*. Given a probability measure μ on $\partial\mathbb{D}$, let $L^2(\partial\mathbb{D}, \mathrm{d}\mu)$ be the space of complex-valued, square-integrable functions with respect to the inner product

$$(2.3) \qquad \langle f, g \rangle_{L^2(\partial\mathbb{D}, \mathrm{d}\mu)} = \int_{\partial\mathbb{D}} \overline{f(e^{i\theta})} g(e^{i\theta}) \, \mathrm{d}\mu(\theta),$$

which we note is antilinear in the first factor. If μ is supported on an infinite number of points, then the functions $\{1, z, z^2, \ldots\}$ are linearly independent in $L^2(\partial\mathbb{D}, \mathrm{d}\mu)$, and hence we can use the Gram–Schmidt procedure to construct the sequence of orthogonal polynomials

$$(2.4) \qquad \Phi_n(z) = \Phi_n(z; \mathrm{d}\mu).$$

That is, $\Phi_n(z; \mathrm{d}\mu)$ is the $L^2(\partial\mathbb{D}, \mathrm{d}\mu)$ projection of z^n on $\{1, z, \ldots, z^{n-1}\}^\perp$. If μ is supported on exactly n distinct points, then the Gram–Schmidt procedure can still be applied but stops with Φ_{n-1}. However, we will assume in the following that this not the case. In particular, our assumptions will ensure that the measures μ that we consider do have infinite support.

Now, note that $\Phi_0(z) = 1$ and the subsequent Φ_n are monic by definition. One of their key properties is that they satisfy the so-called *Szegő recurrence*:

$$(2.5) \qquad \Phi_{n+1}(z) = z\Phi_n(z) - \overline{\alpha}_n \Phi_n^*(z),$$

where the $*$ operation takes the coefficients of an nth degree polynomial, reverses them and then conjugates them, that is,

$$(2.6) \qquad Q_n(z) = \sum_{j=0}^{n} q_j z^j \quad \Longrightarrow \quad Q_n^*(z) = \sum_{j=0}^{n} \overline{q}_{n-j} z^j.$$

The α_n terms that appear in (2.5) are precisely the Verblunsky coefficients, in the sense that μ is the spectral measure for the CMV matrix $\mathcal{C}(\alpha_0, \alpha_1, \ldots)$ and the cyclic vector δ_0, see Theorem 4.2.8 in [39]. Hence the theory of orthogonal polynomials provides a way to realize a probability measure on $\partial\mathbb{D}$ as the spectral measure of a CMV matrix. It also provides a way

to compute the CMV representation of a general unitary operator, by first going through the spectral measure and then pulling out the Verblunsky coefficients.

Another way to reconstruct the probability measure from the orthogonal polynomials is via the *Bernstein–Szegő approximation*. Let φ_n be the normalized orthogonal polynomials defined by

$$\varphi_n := \frac{\Phi_n}{\|\Phi_n\|},$$

where the bars in the denominator indicate the $L^2(\partial\mathbb{D}, \mathrm{d}\mu)$ norm. Since we will use the Bernstein–Szegő approximation repeatedly we record it as a proposition.

PROPOSITION 2.2 (Bernstein–Szegő approximation). *For any sequence $\boldsymbol{\alpha} \in \mathbb{D}^\infty$ the measures*

(2.7) $$\left|\varphi_n(e^{i\theta})\right|^{-2} \frac{\mathrm{d}\theta}{2\pi}$$

are probability measures on $\partial\mathbb{D}$, and they converge weakly to the spectral measure μ as $n \to \infty$.

Although we will not use this fact, it is worth noting that the measures (2.7) also turn out to be the spectral measures for the CMV matrices $\mathcal{C}(\alpha_0, \alpha_1, \ldots, \alpha_{n-1}, 0, 0, 0, \ldots)$.

2.4. *Rotations of the measure.* For a sequence of Verblunsky coefficients $\boldsymbol{\alpha}$ and the corresponding measure $\mu_{\boldsymbol{\alpha}}$, the following result identifies the Verblunsky coefficients of the rotated measure $\mu_{\boldsymbol{\alpha}}(\lambda\cdot)$ for $\lambda \in \partial\mathbb{D}$.

LEMMA 2.3. *For any sequence of Verblunsky $\boldsymbol{\alpha} = \{\alpha_n\}_{n=0}^\infty$ in $\mathbb{D}$, and any $\lambda \in \partial\mathbb{D}$, we have*

$$\Phi_n(\lambda z; \{\alpha_n\}) = \lambda^n \Phi_n(z, \{\lambda^{n+1}\alpha_n\}),$$

for all $z \in \overline{\mathbb{D}}$ and $n \in \mathbb{N}$. In particular, the Verblunsky coefficients of the rotated measure $\mu_{\boldsymbol{\alpha}}(\lambda\cdot)$ are $\{\lambda^{n+1}\alpha_n\}$.

PROOF. The first formula follows by induction on n, by using the Szegő recursion (2.5). Then, the OPUC for the rotated measure are

$$\Phi_n(z; \mu_{\boldsymbol{\alpha}}(\lambda\cdot)) = \overline{\lambda}^n \Phi_n(\lambda z; \mu_{\boldsymbol{\alpha}}),$$

since the right hand side defines a family of monic polynomials which are orthogonal in $L^2(\partial\mathbb{D}, \mu_{\boldsymbol{\alpha}}(\lambda\cdot))$, as can be seen by a simple change of variable. Putting the pieces together, we thus get

$$\Phi_n(z; \mu_{\boldsymbol{\alpha}}(\lambda\cdot)) = \overline{\lambda}^n \Phi_n(\lambda z; \mu_{\boldsymbol{\alpha}}) = \overline{\lambda}^n \Phi_n(\lambda z; \{\alpha_n\}) = \Phi_n(z, \{\lambda^{n+1}\alpha_n\}),$$

which exactly means that the Verblunsky coefficients of $\mu_{\boldsymbol{\alpha}}(\lambda\cdot)$ are $\{\lambda^{n+1}\alpha_n\}$. $\square$

2.5. *Second kind polynomials and the transfer matrix.* The second kind polynomials correspond to rotated versions of the Verblunsky coefficients. For Verblunsky coefficients $\boldsymbol{\alpha} = \{\alpha_n, n \geq 0\}$, the *Aleksandrov measures* are the probability measures on $\partial\mathbb{D}$ with Verblunsky coefficients $\lambda\boldsymbol{\alpha} = \{\lambda\alpha_n, n \geq 0\}$, that is, $\mu_{\lambda\boldsymbol{\alpha}}$, for $\lambda \in \partial\mathbb{D}$. It is important to note that the Aleksandrov measures are *not* just a rotation of the original measure, that is, $\mu_{\lambda\boldsymbol{\alpha}}$ is different from $\mu_{\boldsymbol{\alpha}}(\lambda\cdot)$.

The case $\lambda = -1$ turns out to be very important, and the corresponding polynomials are denoted by Ψ_n and ψ_n, that is,

(2.8) $$\Psi_n(z) = \Phi_n(z; d\mu_{-\alpha}), \qquad \psi_n(z) = \varphi_n(z; d\mu_{-\alpha}).$$

The Ψ_n are referred to as the *second kind polynomials*. The Szegő recursion for Ψ_n becomes

(2.9) $$\Psi_{n+1}(z) = z\Psi_n(z) + \overline{\alpha_n}\Psi_n^*(z).$$

There is a similar Szegő recurrence for (Φ_n^*) and (Ψ_n^*), and their renormalized versions. If we define

$$P_n(z) = \begin{pmatrix} \psi_n(z) & \varphi_n(z) \\ -\psi_n^*(z) & \varphi_n^*(z) \end{pmatrix}$$

then the recursion for $P_n(z)$ is

(2.10) $$P_{n+1}(z) = A_n Z P_n(z), \qquad A_n = \rho_n^{-1/2}\begin{pmatrix} 1 & -\overline{\alpha_n} \\ -\alpha_n & 1 \end{pmatrix}, \qquad Z = \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix}.$$

The matrices A_n (or $A_n Z$) are referred to as the transfer matrices. The recurrence formula for φ_n is particularly useful for our purposes. We write it using an auxiliary quantity $B_n(z)$ defined as

(2.11) $$B_n(z) = \frac{z\varphi_n(z)}{\varphi_n^*(z)}.$$

From (2.10) we can write the Szegő recursion for φ_n as

(2.12) $$\varphi_{n+1}(z) = \rho_n^{-1}\big(z\varphi_n(z) - \overline{\alpha_n}\varphi_n^*(z)\big) = \rho_n^{-1}z\varphi_n(z)\big(1 - \overline{\alpha_n}B_n(z)\big).$$

We will be mostly interested in the case $|z| = 1$, for which the definition of the $*$ operation implies that $\varphi_n^*(e^{i\theta}) = e^{in\theta}\overline{\varphi_n(e^{i\theta})}$. Therefore the definition of B_n becomes

(2.13) $$B_n(e^{i\theta}) = e^{-i(n-1)\theta}\varphi_n(e^{i\theta})/\overline{\varphi_n(e^{i\theta})}$$

from which one easily sees that $B_n(e^{i\theta})$ merely registers the argument of $\varphi_n(e^{i\theta})$ and $|B_n(e^{i\theta})| = 1$. More generally, the definition of φ_n^* implies that B_n is a finite Blaschke product from which the last fact readily follows. From the recursion (2.12) and the definition (2.13) we obtain the recurrence

(2.14) $$B_{n+1}(e^{i\theta}) = e^{i\theta}B_n(e^{i\theta})\frac{1 - \overline{\alpha_n B_n(e^{i\theta})}}{1 - \alpha_n B_n(e^{i\theta})}, \qquad B_0(e^{i\theta}) = e^{i\theta}.$$

The initial condition is from (2.13) and $\varphi_n(e^{i\theta}) = 1$.

Finally, in view of the Berstein–Szegő approximation, we shall need to study the sequence $(|\varphi_n(e^{i\theta})|^{-2})$. Taking the modulus of both sides of (2.12) and using the definition (2.1) of ρ_n gives

(2.15) $$|\varphi_{n+1}(e^{i\theta})|^{-2} = |\varphi_n(e^{i\theta})|^{-2}\frac{1 - |\alpha_n|^2}{|1 - \alpha_n B_n(e^{i\theta})|^2}.$$

This recursion will be very important in our later analysis.

2.6. *Prüfer phases and the modified Verblunsky coefficients.* The recursion formula (2.15) and the Bernstein–Szegő approximation demonstrate that the quantities $\alpha_n B_n(e^{i\theta})$ will play a central role in our study. Due to rotation invariance of the laws of our Verblunsky coefficients, it is usually enough to study φ_n at a single point, for which we use $z = 1$. Following [31], Lemma 2.1 (and similar constructions in [5, 12]) we therefore define the modified Verblunsky coefficients by

$$(2.16) \qquad \gamma_k = \alpha_k B_k(1), \quad k \geq 0.$$

This definition sets up a bijection between the modified Verblunskies and the original ones so we may choose to work with whichever is most convenient at any given moment.

In order to study the field $|\varphi_n(e^{i\theta})|^{-2}$ we will still need to keep track of the B_n field at other points. Given that $B_n(1)$ is already absorbed into the modified Verblunskies, it is enough to keep track of the ratio $B_n(e^{i\theta})/B_n(1)$. Note that this is an evolution on $\partial\mathbb{D}$ so we may define a sequence of continuous, increasing functions $\delta_n : (-\pi, \pi) \to \mathbb{R}$ such that

$$e^{i\delta_n(\theta)} = \frac{B_n(e^{i\theta})}{B_n(1)}.$$

This quantity is known as the *relative Prüfer phase* at time n. Strictly speaking this only defines the process modulo $2\pi\mathbb{Z}$, so the more precise definition we shall use is that δ_n is defined by the recurrence relation

$$(2.17) \qquad \delta_{n+1}(\theta) = \delta_n(\theta) + \theta + 2\Im \log\left(\frac{1 - \gamma_n}{1 - \gamma_n e^{i\delta_n(\theta)}}\right), \quad \delta_0(\theta) = \theta.$$

This recursion is a straightforward consequence of (2.14). The branch of the logarithm is chosen so as to give 0 when $\gamma_n = 0$; see [31], Proposition 2.2, for more details. Note that, by its definition, δ_n is a measurable function of $\gamma_0, \ldots, \gamma_{n-1}$, and that $\delta_n(0) = 0$ for all n. These definitions lead to the commonly used relationship

$$(2.18) \qquad \alpha_k B_k(e^{i\theta}) = \gamma_k e^{i\delta_k(\theta)}.$$

2.7. *Caratheodory functions and the Poisson kernel.* The Poisson kernel for the disk is useful for directly studying μ and for the recurrence relation (2.15). We denote the disk Poisson kernel by $P_\mathbb{D}$, and recall that it is given by

$$(2.19) \qquad P_\mathbb{D}(z, e^{i\theta}) = \frac{1 - |z|^2}{|e^{i\theta} - z|^2}$$

for $z \in \mathbb{D}$. Clearly $P_\mathbb{D}(z, e^{i\theta}) = P_\mathbb{D}(ze^{-i\theta}, 1)$ and recall that $z \mapsto P_\mathbb{D}(z, 1)$ is harmonic since

$$(2.20) \qquad P_\mathbb{D}(z, 1) = \Re\left(\frac{1 + z}{1 - z}\right).$$

For $z_0 \in \mathbb{D}$, let $\phi_{z_0} : \mathbb{D} \to \mathbb{D}$ be the conformal map

$$(2.21) \qquad \phi_{z_0}(z) = \frac{z - z_0}{1 - \overline{z_0}z}.$$

Recall that this is the unique conformal map of the disk to itself which sends z_0 to 0 and has a positive derivative at the origin. It is straightforward to check the relation

$$(2.22) \qquad |\phi'_{z_0}(e^{i\theta})| = P_\mathbb{D}(z_0, e^{i\theta}).$$

Combining the recursion (2.15) and (2.18) with these properties of the Poisson kernels gives

$$(2.23) \qquad |\varphi_n(e^{i\theta})|^{-2} = \prod_{k=0}^{n-1} P_\mathbb{D}(\alpha_k, \overline{B_k(e^{i\theta})}) = \prod_{k=0}^{n-1} P_\mathbb{D}(\gamma_k e^{i\delta_k(\theta)}, 1).$$

We will often take $\theta = 0$ so that the terms on the right hand side are simply $P_{\mathbb{D}}(\gamma_k, 1)$.

The formulas above are useful when the Bernstein–Szegő approximation is used to approximate μ, but the Poisson kernel can also be used to study μ by considering the Caratheodory function:

$$F(z) = F(z; \mathrm{d}\mu) := \int_{\partial\mathbb{D}} \frac{e^{i\theta} + z}{e^{i\theta} - z} \, \mathrm{d}\mu(\theta).$$

Note that the real part of the integrand is the Poisson kernel. The behavior of the modulus $|F(re^{i\theta})|$ as $r \to 1$ gives information on the behavior of the measure, and its magnitude can be bounded in terms of orthogonal polynomials using the *Jitomirskaya–Last inequalities*.

THEOREM 2.4 (Jitomirskaya–Last inequalities [25, 40]). *There exists a universal constant $A > 1$ (independent of the choice of probability measure μ on $\partial\mathbb{D}$) such that*

$$A^{-1} \frac{\|\varphi.(e^{i\theta})\|_{x(r)}}{\|\psi.(e^{i\theta})\|_{x(r)}} \leq |F(re^{i\theta})| \leq A \frac{\|\varphi.(e^{i\theta})\|_{x(r)}}{\|\psi.(e^{i\theta})\|_{x(r)}}$$

for any $r \in [0, 1)$ and $e^{i\theta} \in \partial\mathbb{D}$. The norm is the following: for a sequence $(a_n)_{n\geq 0}$ in $\mathbb{C}$ and $x > 0$ we define $\|a.\|_x$ by

$$\|a.\|_x^2 = \sum_{n=0}^{\lfloor x \rfloor} |a_n|^2 + (x - \lfloor x \rfloor)|a_{\lceil x \rceil}|^2.$$

Then for $r \in [0, 1)$ the function $x(r)$ is the unique solution to

$$(1 - r)\|\varphi.(e^{i\theta})\|_{x(r)}\|\psi.(e^{i\theta})\|_{x(r)} = \sqrt{2}.$$

It is known that at least one of the sequences $(\varphi_n(e^{i\theta}))$ or $(\psi_n(e^{i\theta}))$ is not in $\ell^2(\mathbb{N})$, hence the function $x(r)$ is well defined. The original proof of these inequalities in [25] was for orthogonal polynomials on the line. The translation to the OPUC case is in [40].

2.8. *Dimension theory.* For singular continuous measures there are many different methods for describing their fractal geometry. We use the very precise technique that is described in [25, 32], which in turn comes from the earlier work of Rogers and Taylor [37, 38]. The Rogers and Taylor theory provides a decomposition of a singular continuous measure with respect to a Hausdorff measure, akin to the standard Lebesgue decomposition of a general measure into components that are atomic, absolutely continuous, and singular continuous.

To understand the theory, briefly recall the definition of the Hausdorff measure: for $d \in [0, 1]$ and $S \subset \partial\mathbb{D}$ define $h^d(S)$ by

$$h^d(S) = \lim_{\delta\downarrow 0} \inf_{\delta-\text{covers}} \sum_{i=1}^{\infty} |E_i|^d,$$

where the infimum is over all covers of S by arcs of length less than δ. When restricted to Borel sets, h^d is a Borel measure. For any such S there is a number $\dim_H S$, called the Hausdorff dimension of S, such that $d > \dim_H S \implies h^d(S) = 0$ and $d < \dim_H S \implies h^d(S) = \infty$. Conversely $h^d(S) = 0 \implies \dim_H S \leq d$ and $h^d(S) = \infty \implies \dim_H S \geq d$. Based on these notions, Rogers and Taylor make the following definitions.

DEFINITION 2.5. For $d \in [0, 1]$, a probability measure μ on $\partial\mathbb{D}$:

1. is d-continuous if $\mu(S) = 0$ for all $S \subset \partial\mathbb{D}$ with $h^d(S) = 0$,
2. is d-singular if there exists $S \subset \partial\mathbb{D}$ such that $\mu(S) = 1$ but $h^d(S) = 0$,

3. has exact dimension d if for every $\epsilon > 0$ it is both $(d - \epsilon)$-continuous and $(d + \epsilon)$-singular.

The exact dimension of a measure is a very useful description of the size of the fractal sets that support it. To compute exact dimensions for a measure, Rogers and Taylor introduce the notion of local dimension.

DEFINITION 2.6. Given a probability measure μ on $\partial\mathbb{D}$ and a parameter $s \in [0, 1]$, define, for each $e^{i\theta} \in \partial\mathbb{D}$,

$$D_{\mu}^{s}(\theta) := \limsup_{\epsilon \downarrow 0} \frac{\mu((\theta - \epsilon, \theta + \epsilon))}{(2\epsilon)^{s}},$$

where $(\theta - \epsilon, \theta + \epsilon)$ refers to the arc of length 2ϵ on $\partial\mathbb{D}$ centered around $e^{i\theta}$. Then for each θ, there clearly exists an $s_0 \in [0, 1]$ such that

$$D_{\mu}^{s}(\theta) = \begin{cases} 0 & \text{if } s < s_0, \\ \infty & \text{if } s > s_0. \end{cases}$$

We will refer to this $s_0 = s_0(\mu, \theta)$ as the *local dimension* of the measure μ at the point θ.

With this definition in hand, Rogers and Taylor produce the following important result.

THEOREM 2.7 (Rogers–Taylor [37, 38]). *If the local dimension for μ is s_0 at μ-a.e. points θ, then μ has exact dimension s_0.*

The final result, which is again due to Jitomirskaya and Last and is a consequence of their inequalities (Theorem 2.4), gives a way of computing the local dimension via orthogonal polynomials. In subsequent sections this will be one of our main tools.

PROPOSITION 2.8 (Jitomirskaya–Last [25, 40]). *For $s \in (0, 1)$ let $t = s/(2 - s)$. Then*

$$s > s_0(\mu, \theta) \quad \Longleftrightarrow \quad D_{\mu}^{s}(\theta) = \infty \quad \Longleftrightarrow \quad \liminf_{x \to \infty} \frac{\|\varphi.(e^{i\theta})\|_x}{\|\psi.(e^{i\theta})\|_x^t} = 0.$$

An important corollary of this fact, which we will use repeatedly, is the following. A more refined version of this argument appears in [40], Corollary 10.8.8.

COROLLARY 2.9. *If, for a probability measure μ on $\partial\mathbb{D}$ and a point $e^{i\theta} \in \partial\mathbb{D}$, there are constants $c, d < 1$ such that*

$$\lim_{n \to \infty} \frac{\log |\varphi_n(e^{i\theta})|^{-2}}{\log n} = c, \qquad \lim_{n \to \infty} \frac{\log |\psi_n(e^{i\theta})|^{-2}}{\log n} = d,$$

then $s_0(\mu, \theta) = 2(1 - c)/(2 - c - d)$.

Note that the results are expressed in terms of the negative second power of the orthogonal polynomials, which turns out to be the more useful quantity for random Verblunsky coefficients.

PROOF. By the assumptions, for any $\epsilon > 0$ there exists an N large such that $n \geq N$ implies

$$n^{c-\epsilon} \leq |\varphi_n(e^{i\theta})|^{-2} \leq n^{c+\epsilon}, \qquad n^{d-\epsilon} \leq |\psi_n(e^{i\theta})|^{-2} \leq n^{d+\epsilon}.$$

Thus by summing and using the definition of the $\| \cdot \|_x$ norms and the fact that $c, d < 1$ (so that the norms grow with x) there are constants C_0, C_1 such that

$$C_0 x^{1-c-\epsilon} \le \|\varphi.(e^{i\theta})\|_x^2 \le C_1 x^{1-c+\epsilon}, \qquad C_0 x^{1-d-\epsilon} \le \|\psi.(e^{i\theta})\|_x^2 \le C_1 x^{1-d+\epsilon}.$$

From these inequalities it is straightforward to check that if $t = s/(2-s)$ then

$$\liminf_{x \to \infty} \frac{\|\varphi.(e^{i\theta})\|_x}{\|\psi.(e^{i\theta})\|_x^t} = \begin{cases} \infty, & s < s_0, \\ 0, & s > s_0, \end{cases}$$

with $s_0 = 2(1-c)/(2-c-d)$. Proposition 2.8 completes the proof. $\square$

3. Rotationally invariant independent Verblunsky coefficients.

3.1. *Construction.* From the last subsection we see that dimensional properties of a probability measure can be derived by analyzing the random variable $D_\mu^s(\theta)$, where θ is a random point on $\partial\mathbb{D}$ distributed according to μ. We are interested in the case where μ itself is random and constructed by making a random choice for the Verblunsky coefficients. Letting Q be a probability measure on $\mathbb{D}^\infty$ and $\boldsymbol{\alpha} = \{\alpha_n, n \ge 0\}$ be a generic element of $\mathbb{D}^\infty$, we are therefore interested in the joint probability measure

$$(3.1) \qquad\qquad d\mathbb{P}(\boldsymbol{\alpha}, \theta) := dQ(\boldsymbol{\alpha}) \, d\mu_{\boldsymbol{\alpha}}(\theta)$$

on $\mathbb{D}^\infty \times \partial\mathbb{D}$, where $\mu_{\boldsymbol{\alpha}}$ is the spectral measure corresponding to the sequence $\boldsymbol{\alpha}$. If the pair $(\boldsymbol{\alpha}, \theta)$ is distributed according to this measure, then by Theorem 2.7 dimensional properties for $\mu_{\boldsymbol{\alpha}}$ can be derived by studying the random variable $D_{\mu_{\boldsymbol{\alpha}}}^s(\theta)$.

It is worthwhile to be slightly pedantic about this construction. Throughout our probability space will be $\Omega = \mathbb{D}^\infty \times \partial\mathbb{D}$ equipped with its Borel σ-algebra $\mathcal{F}$ (under the product topology). By the Bernstein–Szegő approximation (2.7), the mapping $\boldsymbol{\alpha} \mapsto \mu_{\boldsymbol{\alpha}}$ is measurable from $\mathbb{D}^\infty$ to the space of probability measures on $\partial\mathbb{D}$ with the weak topology. In other words, μ is a random measure in the usual sense' see [26]. Then, given a probability measure Q on $\mathbb{D}^\infty$, the probability measure $\mathbb{P}$ is defined on rectangles by

$$\mathbb{P}(A \times B) = \int_A \mu_{\boldsymbol{\alpha}}(B) \, dQ(\boldsymbol{\alpha}) = \mathbb{E}_Q\big(\mu_{\boldsymbol{\alpha}}(B) \mathbf{1}_{\{\boldsymbol{\alpha} \in A\}}\big)$$

for Borel sets $A \subset \mathbb{D}^\infty$, $B \subset \partial\mathbb{D}$. Finally $\mathbb{P}$ is extended to all of $\mathcal{F}$ in the usual way. Since the spaces $\mathbb{D}^\infty$ and $\partial\mathbb{D}$ are Polish there are regular conditional distributions for both the $\boldsymbol{\alpha}$ and θ variables; see [27], Theorem 5.3, for a proof of this fact. By the construction of $\mathbb{P}$ the conditional distribution of θ given $\boldsymbol{\alpha}$ is exactly $\mu_{\boldsymbol{\alpha}}$, while the marginal distribution of $\boldsymbol{\alpha}$ is simply Q. In the next section we will analyze the conditional distribution Q_θ of the Verblunsky coefficients $\boldsymbol{\alpha}$ given θ. As pointed out by an anonymous referee, in some literature (see, e.g., [26], Chapter 6), the measure $\mathbb{P}$ is called the Campbell measure, while the conditional measures Q_θ are called the Palm measures. We will also make heavy use of the measure

$$(3.2) \qquad\qquad d\mathbf{P}(\boldsymbol{\alpha}, \theta) = dQ(\boldsymbol{\alpha}) \frac{d\theta}{2\pi},$$

under which the Verblunsky coefficients and θ are independent.

Note that under the joint measure $\mathbb{P}$, we have the following straightforward extension of the Rogers–Taylor result, Theorem 2.7.

COROLLARY 3.1. *If the local dimension for $\mu_{\boldsymbol{\alpha}}$ is s_0 with $\mathbb{P}$-probability one, then $\mu_{\boldsymbol{\alpha}}$ has exact dimension s_0 for Q-a.e. realizations of the Verblunsky coefficients $\boldsymbol{\alpha}$.*

PROOF. A standard application of Fubini's theorem. $\square$

3.2. *Rotationally invariant independent modified Verbunsky coefficients.* From this point on, we enforce Assumption 1.1, namely that the Verblunsky coefficients (α_k) are independent and with a rotationally symmetric distribution. The first easy but extremely useful result concerns the modified Verblunsky coefficients (2.16). Recall that they are defined in Section 2.6 and allow to simplify many formulas. Under Assumption 1.1, using the modified instead of the original coefficients is very convenient because of the following.

LEMMA 3.2 ([31], Lemma 2.1). *The sequence of modified Verblunsky coefficients* $(\gamma_k, k \geq 0)$ *has the same law as the original sequence of Verblunsky coefficients* $(\alpha_k, k \geq 0)$.

This result may seem slightly counterintuitive at first since $B_k(1)$ is itself a function of the α_k, but the important properties is that $B_k(1)$ is a point on $\partial \mathbb{D}$ that is measurable with respect to $\alpha_0, \ldots, \alpha_{k-1}$, and hence, conditionally on $\alpha_0, \ldots, \alpha_{k-1}$, the quantity γ_k is a fixed rotation of α_k and hence has the same law, independent of the actual value of $\alpha_0, \ldots, \alpha_{k-1}$. As a consequence we may choose to work with the modified Verblunsky coefficients instead of the original ones without changing the most important results. We will do so in most of the rest of the paper, although we will freely move between them at many different points.

3.3. *Martingale properties of orthogonal polynomials.* In this section we describe some particular martingales that arise from orthogonal polynomials and their relation to the Bernstein–Szegő approximation.

LEMMA 3.3. *For each fixed $\lambda \in \partial \mathbb{D}$ the measure $\mathbb{P}$ is invariant under the transformation*

$$\big(\{\alpha_n\}_{n=0}^{\infty}, e^{i\theta}\big) \mapsto \big(\{\lambda^{n+1}\alpha_n\}_{n=0}^{\infty}, \overline{\lambda}e^{i\theta}\big).$$

PROOF. By assumption, the sequences $\{\alpha_n, n \geq 0\}$ and $\{\lambda^{n+1}\alpha_n, n \geq 0\}$ have the same law, hence so too do the associated random measures. By Lemma 2.3 these measures are $\mu_{\boldsymbol{\alpha}}(\cdot)$ and $\mu_{\boldsymbol{\alpha}}(\lambda\cdot)$, hence the result follows. $\square$

PROPOSITION 3.4. *Under the measure* $\mathbf{P}$, *the process* $n \mapsto |\varphi_n(e^{i\theta})|^{-2}$ *is a martingale with respect to the filtration* $\mathcal{F}_{n-1}^{\theta} := \sigma(\theta, \alpha_0, \alpha_1, \ldots, \alpha_{n-1})$.

The martingale $|\varphi_n(e^{i\theta})|^{-2}$ also appears in a crucial way in [13].

PROOF. First observe that by the recursions (2.23) and (2.17) the variables $\varphi_n(e^{i\theta})$ and $\delta_n(\theta)$ are measurable with respect to $\mathcal{F}_{n-1}^{\theta}$ (the shift in the indexing is because φ_0 and δ_0 are deterministic by convention, while φ_1 and δ_1 are determined by the Verblunsky coefficient α_0, and so on). Using this measurability and recursion (2.23) we have the relation

$$\mathbb{E}_{\mathbf{P}}\big[\big|\varphi_{n+1}(e^{i\theta})\big|^{-2}|\mathcal{F}_{n-1}^{\theta}\big] = \big|\varphi_n(e^{i\theta})\big|^{-2}\mathbb{E}_{\mathbf{P}}\big[P_{\mathbb{D}}(\gamma_n e^{i\delta_n(\theta)}, 1)|\mathcal{F}_{n-1}^{\theta}\big].$$

In the remaining expectation the term $\delta_n(\theta)$ is $\mathcal{F}_{n-1}^{\theta}$-measurable, whereas γ_n is independent of $\mathcal{F}_{n-1}^{\theta}$. Recall as well that γ_n has a rotationally invariant law, by Lemma 3.2. Therefore

$$\mathbb{E}_{\mathbf{P}}\big[P_{\mathbb{D}}(\gamma_n e^{i\delta_n(\theta)}, 1)|\mathcal{F}_{n-1}^{\theta}\big] = \mathbb{E}_{\mathbf{P}}\big[P_{\mathbb{D}}(\gamma_n, 1)\big] = \mathbb{E}_Q\big[P_{\mathbb{D}}(0, 1)\big] = 1.$$

The second-to-last equality uses that γ_n is rotation invariant and that $z \mapsto P_{\mathbb{D}}(z, 1)$ is harmonic, so that its integral on a circle is just the value at the center, that is, $P_{\mathbb{D}}(0, 1) = 1$. $\square$

REMARK 3.5. To be pedantic again, the filtration $\mathcal{F}_n^\theta$ is the Borel σ-algebra generated by the θ variable and the first $n+1$ elements of a sequence in $\mathbb{D}^\infty$, that is,

$$\mathcal{F}_n^\theta = \sigma\left(\underset{i=0}{\overset{n}{\times}} \mathcal{B}(\mathbb{D}) \times \underset{i=n+1}{\overset{\infty}{\times}} \{\varnothing, \mathbb{D}\} \times \mathcal{B}(\partial\mathbb{D}) \right),$$

where $\mathcal{B}(\mathbb{D})$ is the Borel σ-algebra for $\mathbb{D}$, $\mathcal{B}(\partial\mathbb{D})$ is the Borel σ-algebra for $\partial\mathbb{D}$, and $\{\varnothing, \mathbb{D}\}$ is the trivial σ-algebra for $\mathbb{D}$. Note that these $\mathcal{F}_n^\theta$ generate the entire σ-algebra for Ω, that is,

$$\mathcal{F} = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n^\theta \right).$$

For consistency of notation we write $\mathcal{F}_{-1}^\theta$ to be the σ-algebra for θ, that is, $\mathcal{F}_{-1}^\theta = \sigma(\times_{i\geq 0}\{\varnothing, \mathbb{D}\} \times \mathcal{B}(\partial\mathbb{D}))$. Forgetting the θ variable entirely we can also consider the smaller σ-algebras

$$\mathcal{F}_n = \sigma\left(\underset{i=0}{\overset{n}{\times}} \mathcal{B}(\mathbb{D}) \times \underset{i=n+1}{\overset{\infty}{\times}} \{\varnothing, \mathbb{D}\} \times \{\varnothing, \partial\mathbb{D}\} \right).$$

Slightly abusing notation, we will also think of $(\mathcal{F}_n)$ as the standard filtration for $\mathbb{D}^\infty$.

REMARK 3.6. There is a slightly weaker version of Proposition 3.4 that at first may seem more natural: for any fixed θ a virtually identical argument shows that $|\varphi_n(e^{i\theta})|^{-2}$ is a martingale with respect to $\mathcal{F}_n$, under the measure Q. This gives the relation

(3.3) $$\mathbb{E}_Q\big[|\varphi_{n+1}(e^{i\theta})|^{-2}|\mathcal{F}_{n-1}\big] = |\varphi_n(e^{i\theta})|^{-2} \quad Q\text{-a.s.}$$

However note that there is a null set involved in this identity, and by fixing θ the null set may depend on θ. The union of these null sets over the uncountably many θ may become an event of positive probability. Proposition 3.4 avoids this by working with respect to the joint measure and gives that there is only one single null set in $\mathbb{D}^\infty \times \partial\mathbb{D}$ on which the martingale relation does not hold. Also note that in the proof of Proposition 3.4 the Lebesgue measure $d\theta/2\pi$ could have been replaced by any fixed probability measure on $\partial\mathbb{D}$. This determines what measure the null set is with respect to, but for our purposes, the Lebesgue measure will be sufficient.

The latter remark is useful for the next result, which is noteworthy because it provides a probabilistic proof of the convergence part of the Bernstein–Szegő approximation. The following argument also appears in [13].

COROLLARY 3.7. *For independent and rotationally invariant Verblunsky coefficients, the probability measures* $|\varphi_n(e^{i\theta})|^{-2}\,d\theta/2\pi$ *Q-almost surely converge* (*weakly*) *as* $n \to \infty$.

PROOF. Let $f : \partial\mathbb{D} \to \mathbb{R}$ be continuous and hence bounded. Using Proposition 3.4, it is straightforward to show that the process

$$n \mapsto \int_{\partial\mathbb{D}} f(e^{i\theta})|\varphi_n(e^{i\theta})|^{-2}\frac{d\theta}{2\pi}$$

is a martingale with respect to Q and $(\mathcal{F}_{n-1}, n \geq 0)$, and since it is bounded (because f is and $|\varphi_n|^{-2}\,d\theta/2\pi$ is a probability measure) the martingale convergence theorem implies that it converges almost surely. Since the almost sure convergence holds simultaneously for a countable and dense collection of f, this implies a.s. weak convergence (see [26]). $\square$

Note that this only provides a probabilistic proof of the convergence part of the Bernstein–Szegő approximation. We do not have a separate proof that $|\varphi_n(e^{i\theta})|^{-2}\,d\theta/2\pi$ are probability measures (which is implicitly used in the above proof), nor are we able to identify the limiting measure as being the spectral measure. However, the Bernstein–Szegő approximation and the independence of the Verblunsky coefficient implies that the conditional expectation of the spectral measure $\mu_{\boldsymbol{\alpha}}$ given $\mathcal{F}_n$ is the measure $|\varphi_{n+1}(e^{i\theta})|^{-2}\,d\theta/2\pi$, in the following sense.

LEMMA 3.8. *Let $F : \mathbb{D}^\infty \times \partial\mathbb{D} \to \mathbb{R}$ be continuous and $\mathcal{F}_n^\theta$-measurable. For $n \geq -1$*

$$\mathbb{E}_{\mathbb{P}}\big[F(\boldsymbol{\alpha},\theta)|\mathcal{F}_n\big] = \mathbb{E}_Q\bigg[\int_{\partial\mathbb{D}} F(\boldsymbol{\alpha},\theta)\,d\mu_{\boldsymbol{\alpha}}(\theta)|\mathcal{F}_n\bigg] = \int_{\partial\mathbb{D}} F(\boldsymbol{\alpha},\theta)|\varphi_{n+1}(e^{i\theta})|^{-2}\frac{d\theta}{2\pi}.$$

In particular, for $A \in \mathcal{F}_n$ and $B \in \mathcal{B}(\partial\mathbb{D})$ we have

$$\mathbb{P}(A \times B) = \int_A \int_B |\varphi_{n+1}(e^{i\theta})|^{-2}\frac{d\theta}{2\pi}\,dQ(\boldsymbol{\alpha}).$$

The proof is very similar to that of Corollary 3.7. Finally we note that there are probabilistic proofs concerning the absolute continuity and singularity of the spectral measure $\mu_{\boldsymbol{\alpha}}$ with respect to Lebesgue measure.

PROPOSITION 3.9. *With respect to the measure $d\mathbf{P}$:*

(i) *the event $\{|\varphi_n(e^{i\theta})|^{-2} \to 0\}$ has either probability 1 or probability 0;*
(ii) *if $|\varphi_n(e^{i\theta})|^{-2} \to 0$ with probability 1, then with Q-probability one, $\mu_{\boldsymbol{\alpha}}$ is singular with respect to Lebesgue measure;*
(iii) *if $|\varphi_n(e^{i\theta})|^{-2}$ converges in L^1, then with Q-probability 1, $\mu_{\boldsymbol{\alpha}}$ is absolutely continuous with respect to the Lebesgue measure.*

PROOF. For ease of notation write $X(\theta) = \limsup_{n\to\infty} |\varphi_n(e^{i\theta})|^{-2}$. By the martingale property of Proposition 3.4 we know that $|\varphi_n(e^{i\theta})|^{-2}$ converges to $\mathbf{P}$-almost surely as $n \to \infty$, and in particular $\mathbf{P}(X < \infty) = 1$. To prove part (i), note that the event $\{X = 0\}$ is a tail event for the filtration $\mathcal{F}_n^\theta$ (by, e.g., equation (2.23) and the positivity of the Poisson kernel), and since $\mathbf{P}$ is a product measure, we have by the Kolmogorov 0-1 law that $\mathbf{P}(X = 0) \in \{0, 1\}$.

For parts (ii) and (iii) first define measures $\mathbb{P}_{ac}$ and $\mathbb{P}_s$ by

$$d\mathbb{P}_{ac} = X\,d\mathbf{P}, \qquad \mathbb{P}_s(A) = \mathbb{P}\big(A \cap \{X = \infty\}\big),$$

for Borel $A \subset \Omega$. Since $\mathbf{P}(X < \infty) = 1$, we have $\mathbb{P}_{ac} \ll \mathbf{P}$ and $\mathbb{P}_s \perp \mathbf{P}$ (hence the reason for the names). By classical arguments [16], Theorem 5.3.3, we have that $\mathbb{P} = \mathbb{P}_{ac} + \mathbb{P}_s$, which is the Lebesgue decomposition of $\mathbb{P}$ with respect to $\mathbf{P}$. In case (ii) we have that $\mathbf{P}(X = 0) = 1$ and thus $\mathbb{P}$ is singular with respect to $\mathbf{P}$. Then, by a straightforward Fubini argument, we see that $\mathrm{Leb}(\{e^{i\theta} \in \partial\mathbb{D} : (\boldsymbol{\alpha}, e^{i\theta}) \in \Omega_0\}) = 0$, Q-a.e., as desired. Part (iii) is handled similarly. □

3.4. *The conditional measure for Verblunsky coefficients.* Now we turn to the task of describing the conditional distribution Q_θ for the joint measure $\mathbb{P}$. The intuition comes from the Bernstein–Szegő approximation, which says that

$$d\mathbb{P}(\boldsymbol{\alpha},\theta) = dQ(\boldsymbol{\alpha})\,d\mu_{\boldsymbol{\alpha}}(\theta) = \lim_{n\to\infty} dQ(\boldsymbol{\alpha})|\varphi_n(e^{i\theta})|^{-2}\,d\theta/2\pi.$$

The next lemma shows that the marginal distribution of θ is Lebesgue, and therefore the formula above suggests that

$$dQ_\theta(\boldsymbol{\alpha}) = \lim_{n\to\infty} |\varphi_n(e^{i\theta})|^{-2}\,dQ(\boldsymbol{\alpha}).$$

A proof of this relation follows after. It consists of showing that the limit above exists and that it agrees with the conditional measure.

LEMMA 3.10. *Under the measure $\mathbb{P}$ the marginal distribution of θ is Lebesgue.*

PROOF 1. Fix $\lambda \in \partial \mathbb{D}$. By Lemma 3.3 the random variables $e^{i\theta}$ and $\lambda e^{i\theta}$ have the same law under $\mathbb{P}$, that is, the law is invariant under all rotations. Therefore it must be Lebesgue. $\square$

PROOF 2. By definition of $\mathbb{P}$ the marginal measure for θ is the probability measure ν on $\partial \mathbb{D}$ with

$$\nu(B) = \mathbb{P}(\mathbb{D}^\infty \times B) = \mathbb{E}_Q[\mu_{\boldsymbol{\alpha}}(B)]$$

for all Borel subsets B of $\partial \mathbb{D}$. However, since $\mathbb{D}^\infty \in \mathcal{F}_{-1}$ the second part of Lemma 3.8 implies

$$\mathbb{P}(\mathbb{D}^\infty \times B) = \int_{\mathbb{D}^\infty} \int_B |\varphi_0(e^{i\theta})|^{-2} \frac{d\theta}{2\pi} \, dQ(\boldsymbol{\alpha}) = \int_B \frac{d\theta}{2\pi}.$$

The last equality follows from $\varphi_0 = 1$. $\square$

Note the first proof is based on independence and rotation invariance of the Verblunsky coefficients while the second uses the martingale property. If the martingale property can be established without assuming independence and rotation invariance then the second proof still goes through.

PROPOSITION 3.11. *For each θ, the limiting measure*

$$dQ_\theta(\boldsymbol{\alpha}) := \lim_{n \to \infty} |\varphi_n(e^{i\theta})|^{-2} \, dQ(\boldsymbol{\alpha})$$

exists, and, under $\mathbb{P}$, is the conditional measure of $\boldsymbol{\alpha}$ given θ.

PROOF. The proof of existence is standard. Let $A \in \mathcal{F}_n$. Then by the martingale property (3.3),

$$\mathbb{E}_Q[\mathbf{1}_{\{\boldsymbol{\alpha} \in A\}} |\varphi_m(e^{i\theta})|^{-2}] = \mathbb{E}_Q[\mathbf{1}_{\{\boldsymbol{\alpha} \in A\}} |\varphi_{n+1}(e^{i\theta})|^{-2}]$$

for all $m \geq n + 1$. Thus the sequence of measures $|\varphi_n(e^{i\theta})|^{-2} \, dQ(\alpha)$, each one defined on $\mathcal{F}_n$, form a consistent family, and therefore, by the Kolmogorov extension theorem [36], Chapter 1, Theorem 3.2, they extend uniquely to a measure Q_θ on $\sigma(\bigcup_{n \geq 0} \mathcal{F}_n) = \mathcal{B}(\mathbb{D}^\infty)$. Furthermore, by the equality above, we have

$$Q_\theta(A) = \lim_{m \to \infty} \mathbb{E}_Q[\mathbf{1}_{\{\boldsymbol{\alpha} \in A\}} |\varphi_m(e^{i\theta})|^{-2}] = \mathbb{E}_Q[\mathbf{1}_{\{\boldsymbol{\alpha} \in A\}} |\varphi_{n+1}(e^{i\theta})|^{-2}].$$

This is sufficient to show that Q_θ is the weak limit of these measures [4], Theorem 2.2.

To show that Q_θ is the conditional distribution of $\boldsymbol{\alpha}$, first observe that the fact that it can be written as a limit implies that $\theta \mapsto Q_\theta(A)$ is measurable for each $A \in \mathcal{B}(\mathbb{D}^\infty)$. Hence it makes sense to consider the measure $\mathbb{P}'$ on Ω defined by

$$\mathbb{P}'(A \times B) = \int_B Q_\theta(A) \frac{d\theta}{2\pi}$$

for $A \in \mathcal{B}(\mathbb{D}^\infty)$, and $B \in \mathcal{B}(\partial \mathbb{D})$. Now if $A \in \mathcal{F}_n$ then Lemma 3.8 and Fubini's theorem ensure that $\mathbb{P}'(A \times B) = \mathbb{P}(A \times B)$, and a π-λ argument [16] ensures that $\mathbb{P} = \mathbb{P}'$. $\square$

From the definition of Q_θ, we can quickly derive some simple properties.

PROPOSITION 3.12. *The following statements hold.*

(i) *If $|\varphi_n(e^{i\theta})|^{-2} \to 0$ with positive Q-probability, then the measures Q and Q_θ are singular.*

(ii) *If $|\varphi_n(e^{i\theta})|^{-2}$ converges in L^1, then the measures Q and Q_θ are absolutely continuous.*

(iii) *For any $\theta \in \partial\mathbb{D}$, if $\{\alpha_n\}_{n=0}^\infty$ has law Q_θ, then $\{e^{i(n+1)\theta}\alpha_n\}_{n=0}^\infty$ has law Q_0.*

(iv) *If $(\{\alpha_n\}_{n=0}^\infty, \theta)$ has law $\mathbb{P}$, then $\{e^{i(n+1)\theta}\alpha_n\}_{n=0}^\infty$ has law Q_0.*

PROOF. For parts (i) and (ii), let $X(\theta) = \limsup_{n\to\infty} |\varphi_n(e^{i\theta})|^{-2}$. For each fixed θ, the fact that $(|\varphi_n(e^{i\theta})|^{-2})$ is a martingale shows that Q-a.s., the limsup above is merely a limit. In particular $Q(X(\theta) < \infty) = 1$. We can then define measures $Q_{\theta,ac}$ and $Q_{\theta,s}$ by

$$dQ_{\theta,ac} = X(\theta)\,dQ, \qquad Q_{\theta,s}(A) = Q_\theta(A \cap \{X(\theta) = \infty\})$$

and conclude as in the proof of Proposition 3.9.

For part (iii), we have that, for any F that is $\mathcal{F}_{n-1}$-measurable,

$$\begin{aligned}
\mathbb{E}_{Q_\theta}[F(\boldsymbol{\alpha})] &= \mathbb{E}_Q\big[|\varphi_n(e^{i\theta}; \boldsymbol{\alpha})|^{-2} F(\boldsymbol{\alpha})\big]\\
&= \mathbb{E}_Q\big[|\varphi_n(e^{i\theta}; \{e^{-i(k+1)\theta}\alpha_k\})|^{-2} F(\{e^{-i(k+1)\theta}\alpha_k\})\big]\\
&= \mathbb{E}_Q\big[|\varphi_n(1; \boldsymbol{\alpha})|^{-2} F(\{e^{-i(k+1)\theta}\alpha_k\})\big] = \mathbb{E}_{Q_0}\big[F(\{e^{-i(k+1)\theta}\alpha_k\})\big].
\end{aligned}$$

The first equality is the definition of Q_θ (see the previous proof), the second holds by rotation invariance of Q, and the third is by Lemma 2.3. Finally, part (iv) follows from (iii) and the decomposition $d\mathbb{P}(\boldsymbol{\alpha}, \theta) = dQ_\theta(\boldsymbol{\alpha})\,d\theta/2\pi$, since for Borel $A \subset \mathbb{D}^\infty$,

$$\mathbb{P}(\{e^{i(n+1)\theta}\alpha_n\}_{n=0}^\infty \in A) = \int_{\partial\mathbb{D}} Q_\theta(\{e^{i(n+1)\theta}\alpha_n\}_{n=0}^\infty \in A)\,d\theta/2\pi = Q_0(\{\alpha_n\}_{n=0}^\infty \in A). \quad\square$$

Parts (iii) and (iv) show that it is enough to study the measures Q_0. In particular, part (iii) combined with Corollary 3.1 gives the following simplification for computing exact dimensions.

COROLLARY 3.13. *Suppose there are constants $c, d < 1$ such that, with Q_0 probability one,*

$$\lim_{n\to\infty} \frac{\log|\varphi_n(1; \boldsymbol{\alpha})|^{-2}}{\log n} = c, \qquad \lim_{n\to\infty} \frac{\log|\psi_n(1; \boldsymbol{\alpha})|^{-2}}{\log n} = d.$$

Then μ_α has exact dimension $2(1-c)/(2-c-d)$ for Q-a.e. realizations of the Verblunsky coefficients.

PROOF. By Corollary 3.1 it is sufficient to show that $s_0(\mu_{\boldsymbol{\alpha}}, \theta) = 2(1-c)/(2-c-d)$ holds $\mathbb{P}$ almost surely. By a rotation of the circle, it is clear that for any sequence $\boldsymbol{\alpha}$ of Verblunsky coefficients there is the identity $s_0(\mu_{\boldsymbol{\alpha}}, \theta) = s_0(\mu_{\boldsymbol{\alpha}}(e^{i\theta}\cdot), 0)$. By Lemma 2.3 we have that

$$\mu_{\boldsymbol{\alpha}}(e^{i\theta}\cdot) = \mu_{\{e^{i(n+1)\theta}\alpha_n\}}.$$

Therefore by part (iv) of Proposition 3.12, one has

$$\mathbb{P}(s_0(\mu_{\boldsymbol{\alpha}}, \theta) \le s) = \mathbb{P}(s_0(\mu_{\boldsymbol{\alpha}}(e^{i\theta}\cdot), 0) \le s) = Q_0(s_0(\mu_{\boldsymbol{\alpha}}, 0) \le s)$$

for arbitrary s, and the result then follows from Corollary 2.9. $\square$

3.5. *The Markov chain description for the conditional Verblunskies.* In this section, we analyze the behavior of the Verblunsky coefficients under the measure Q_0. The description turns out to be simplest in terms of the modified Verblunsky coefficients γ_n, which maintain their independence properties but lose their rotation invariance. The radial part of each Verblunsky maintains the same law but its angular part is biased towards $z = 1$. Translating these results back into the original Verblunsky coefficients shows that the quantity $(\alpha_n, B_n(1))$ forms a Markov chain with a straightforward transition probability.

THEOREM 3.14.   *Under Q_0 the modified Verblunsky coefficients $\{\gamma_n : n \geq 0\}$ are mutually independent with marginal laws*

$$(3.4) \qquad Q_0(\gamma_n \in dz) = P_{\mathbb{D}}(z, 1) Q(\gamma_n \in dz).$$

Consequently, under Q_0 the process $n \mapsto (\alpha_n, B_n(1))$ forms a Markov chain with transition kernel

$$(3.5) \qquad Q_0(\alpha_{n+1} \in dz | \alpha_n, B_n(1)) = P_{\mathbb{D}}(\alpha_n B_n(1), 1) Q(\alpha_{n+1} \in dz),$$

while $B_n(1)$ updates by formula (2.14).

PROOF.   This is a straightforward consequence of the independence and rotation invariance of the modified Verblunsky coefficients under Q and formula (2.23) at $\theta = 0$, which shows that

$$|\varphi_n(1)|^{-2} = \prod_{k=0}^{n-1} P_{\mathbb{D}}(\gamma_k, 1),$$

so that the Radon–Nikodym derivative between Q_0 and Q maintains the product structure of Q. The Markov chain description for the original Verblunsky coefficients follows from (2.15) expressed in terms of α_n and $B_n(1)$, or alternatively by taking the result for γ_n and applying the bijection from modified Verblunkies back to the original ones.   □

3.6. *Coupling Q and Q_0.* In this section we derive an algorithm that converts Verblunskys distributed according to Q into Verblunskys distributed according to Q_0. We use it to derive several distributional properties of the Verblunskys under Q_0.

PROPOSITION 3.15.   *Given a sequence $\{\gamma_n\}_{n=0}^{\infty} \in \mathbb{D}^{\infty}$ define a new sequence $\{\gamma_n^*\}_{n=0}^{\infty}$ by*

$$(3.6) \qquad \gamma_n^* := \gamma_n \frac{1 + \overline{\gamma_n}}{1 + \gamma_n}.$$

If $\{\gamma_n\}_{n=0}^{\infty}$ has law Q, then $\{\gamma_n^*\}_{n=0}^{\infty}$ has law Q_0. Consequently, under Q_0 the modified Verblunsky coefficients are invariant under conjugation, have radial laws that are the same as they are under Q, and have integer moments given by

$$(3.7) \qquad \mathbb{E}_{Q_0}[\gamma_n^m] = \mathbb{E}_Q[(\gamma_n^*)^m] = \mathbb{E}_Q[|\gamma_n|^{2m}].$$

In the above we of course think of the γ_n as the modified Verblunksy coefficients. It is straightforward to modify this result to obtain the analogous result for the original Verblunsky coefficients, if so desired. The proof relies on the following simple lemma.

LEMMA 3.16.   *Let Z be a random variable taking values in $\mathbb{D}$ whose law $d\Lambda(z)$ is rotationally invariant. Then, for fixed $w \in \partial \mathbb{D}$, the random variable*

$$|Z| \phi_{-|Z|w}\left(\frac{Z}{|Z|}\right) = Z \frac{1 + \overline{Z}w}{1 + Z\overline{w}}$$

has law $P_{\mathbb{D}}(z, w) d\Lambda(z)$ on $\mathbb{D}$.

PROOF. Represent Z by its radial and angular parts via the pair $(|Z|, Z/|Z|)$. The rotation invariance of Z implies that the joint measure of this pair can be written as $d\eta(r)\,d\theta/2\pi$, where η is the distribution of $|Z|$. Then recall that for $z_0 \in \mathbb{D}$, the conformal map ϕ_{z_0} sends the disk to itself and z_0 to 0, and on the circle has derivative

$$|\phi'_{z_0}(e^{i\theta})| = P_{\mathbb{D}}(z_0, e^{i\theta}).$$

Therefore if X is a uniform random variable on $\partial\mathbb{D}$, then $\phi_{-z_0}(X)$ has law

$$|\phi'_{z_0}(e^{i\theta})|\frac{d\theta}{2\pi}.$$

Conditioning on $|Z|$ and then using the identity

$$|\phi'_{rw}(e^{i\theta})| = P_{\mathbb{D}}(rw, e^{i\theta}) = P_{\mathbb{D}}(r\overline{w}, e^{-i\theta}) = P_{\mathbb{D}}(re^{i\theta}, w) = |\phi'_{re^{i\theta}}(w)|$$

we get that $\phi_{-|Z|w}(Z/|Z|)$ has law

$$|\phi'_{|Z|w}(e^{i\theta})|\frac{d\theta}{2\pi} = |\phi'_{|Z|e^{i\theta}}(w)|\frac{d\theta}{2\pi}.$$

Multiplying by the marginal density of the $|Z|$ variable completes the proof. $\square$

PROOF OF PROPOSITION 3.15. Lemma 3.16 and Theorem 3.14 combine to show that the law of γ_n^* under Q is exactly Q_0. Conjugation invariance of γ_n under Q (which follows from rotation invariance) therefore implies conjugation invariance of γ_n^*, by (3.6). That the radial laws are the same under Q and Q_0 follows from $|\gamma_n| = |\gamma_n^*|$. Finally, for the moment formula, use that

$$\mathbb{E}_{Q_0}[\gamma_n^m] = \mathbb{E}_Q[(\gamma_n^*)^m] = \mathbb{E}_Q\left[\left(\gamma_n\frac{1+\overline{\gamma_n}}{1+\gamma_n}\right)^m\right] = \mathbb{E}_Q\left[|\gamma_n|^m\phi_{-|\gamma_n|}\left(\frac{\gamma_n}{|\gamma_n|}\right)^m\right].$$

Since the function $z \mapsto \phi_{-|\gamma_n|}(z)^m$ is analytic on $\overline{\mathbb{D}}$ and $\gamma_n/|\gamma_n|$ is uniformly distributed on $\partial\mathbb{D}$ and independent of $|\gamma_n|$ it follows that

$$\mathbb{E}_Q\left[\phi_{-|\gamma_n|}\left(\frac{\gamma_n}{|\gamma_n|}\right)^m \Big| |\gamma_n|\right] = \phi_{-|\gamma_n|}(0)^m = |\gamma_n|^m.$$

The tower property of conditional expectation finishes the computation. $\square$

4. Almost sure local dimensions.

4.1. *Main result.* In this section we compute exact local dimensions for μ_α at typical points of the measure. The techniques we use are very general and apply to *any* choice of independent and rotationally invariant Verblunsky coefficients satisfying (1.5) and (1.6), to wit

(4.1) $$\mathbb{E}_Q[|\alpha_n|^2] \sim \frac{2}{\beta n}$$

and

(4.2) $$\mathbb{E}_Q[|\alpha_n|^3] = o\left(\frac{1}{n}\right).$$

We shall prove Theorem 1.4, namely that Q-a.s., μ_α has exact Hausdorff dimension $1 - 2/\beta$. By Corollaries 2.9 and 3.13 it is enough to prove the following.

PROPOSITION 4.1. *With Q_0-probability one*

$$\lim_{n\to\infty}\frac{\log|\varphi_n(1)|^{-2}}{\log n} = \frac{2}{\beta}, \qquad \lim_{n\to\infty}\frac{\log|\psi_n(1)|^{-2}}{\log n} = -\frac{2}{\beta}.$$

Note that the assumption of $\beta > 2$ is only used in this section to satisfy the assumption of Corollary 2.9 that $c < 1$. All subsequent results hold for all $\beta > 0$.

4.2. *Simplification to bounded Verblunsky coefficients.* In all the proofs, we will in fact assume that the α_n are uniformly bounded away from 1, that is, for some $\delta > 0$,

(4.3)
$$Q(\forall n \geq 0 \, |\alpha_n| \leq 1 - \delta) = 1.$$

Here is why this is enough to obtain the results for general variables α_n satisfying (4.1) and (4.2). First note that (4.1) implies that the sequence $|\alpha_n|$ converges to zero in probability. Since it also take values in $\mathbb{D}$, for any $\varepsilon > 0$ we may find $\delta > 0$ such that

$$Q(\forall n \geq 0 \, |\alpha_n| \leq 1 - \delta) \geq 1 - \varepsilon.$$

Then, truncate the variables at radius $1 - \delta$ by letting

$$\widetilde{\alpha}_n = (|\alpha_n| \wedge 1 - \delta) \frac{\alpha_n}{|\alpha_n|}.$$

Our choice of δ means that

$$Q(\forall n \geq 0 \, \alpha_n = \widetilde{\alpha}_n) \geq 1 - \varepsilon.$$

The variables $(\widetilde{\alpha}_n)$ are still rotationally invariant and independent and satisfy (4.3) and (4.2). Moreover, by Hölder's and Markov's inequalities,

$$\mathbb{E}_Q\big[|\alpha_n|^2 \mathbf{1}_{\{|\alpha_n|>1-\delta\}}\big] \leq \mathbb{E}\big[|\alpha_n|^3\big]^{2/3} \mathbb{P}(|\alpha_n| > 1 - \delta)^{1/3}$$
$$\leq \mathbb{E}\big[|\alpha_n|^3\big]^{2/3} \mathbb{E}\big[|\alpha_n|^2\big]^{1/3} (1 - \delta)^{-2/3}$$
$$= o(n^{-2/3})n^{-1/3} = o(n^{-1}),$$

so that

$$\mathbb{E}_Q[|\widetilde{\alpha}_n|^2] \geq \mathbb{E}_Q[|\alpha_n|^2 \mathbf{1}|\alpha_n| \leq 1 - \delta] = \mathbb{E}_Q[|\alpha_n|^2] - \mathbb{E}_Q[|\alpha_n|^2 \mathbf{1}|\alpha_n| > 1 - \delta] = \frac{2}{\beta n} + o\Big(\frac{1}{n}\Big).$$

Combining this with the obvious inequality $\mathbb{E}_Q[|\widetilde{\alpha}_n|^2] \leq \mathbb{E}_Q[|\alpha_n|^2]$ we have

$$\mathbb{E}_Q(|\widetilde{\alpha}_n|^2) \sim \frac{2}{\beta n}.$$

So the second moments of the $\widetilde{\alpha}_n$ and of α_n have the same asymptotic rate of decay, with the same constant. So once we prove, for instance, that $\mu_{\widetilde{\alpha}}$ has a.s. exact Hausdorff dimension $1 - 2/\beta$, then μ_α has exact Hausdorff dimension $1 - 2/\beta$ with probability at least $1 - \varepsilon$. Since this holds for all $\varepsilon > 0$, and the result does not depend on ε, then this also proves it for μ_α.

The main purpose of assuming (4.3) is that it allows us to make Taylor expansions of functions of α_n that are uniform in $\boldsymbol{\alpha}$. For instance, we have the Taylor series expansion

(4.4)
$$\log P_{\mathbb{D}}(z, 1) = 2\Re z - 2(\Im z)^2 + O(|z|^3),$$

when $z \to 0$. Applied to formula (2.23) for $|\varphi_n(1)|^{-2}$, this gives

(4.5)
$$\log |\varphi_n(1)|^{-2} = \sum_{k=0}^{n-1} (2\Re \gamma_k - 2(\Im \gamma_k)^2 + O(|\gamma_k|^3)).$$

The moment assumptions imply that the first two terms grow like a multiple of $\log n$ (see below), but the last one is more tricky to deal with. For instance, without knowing more, we cannot take an expectation of this formula. More precisely, we cannot put the expectation inside the $O(\cdot)$. But if the α_n do satisfy (4.3), then $O(\cdot)$ simply means that there is a constant C, that depends only on δ, such that

$$O(|\alpha_k|^3) \leq C |\alpha_k|^3.$$

Therefore

$$\mathbb{E}\left[\sum_{k=0}^{n-1} O(|\alpha_k|^3)\right] \leq C \sum_{k=0}^{n-1} \mathbb{E}(|\alpha_k|^3) = o(\log n),$$

by (4.2). As a conclusion, these terms will not matter.

4.3. *A lemma on martingales.* For the precise asymptotics of the first two terms in the Taylor expansion above it will be enough to analyze their means, as a result of the following result, which is taken from [43], Chapter 12.

LEMMA 4.2. *Let M_n be a martingale with $M_0 = 0$ and with increments $X_n = M_n - M_{n-1}$ that satisfy $\mathbb{E}[X_n^2] < \infty$ for all n. Assume that (b_n) is a sequence of positive real numbers s.t.*

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}(X_n^2)}{b_n^2} < \infty.$$

Then $M_n/b_n \to 0$ almost surely as $n \to \infty$.

4.4. *Asymptotics under Q.* Before proving Proposition 4.1, we quickly analyze the asymptotics of $|\varphi_n(1)|^{-2}$ and $|\psi_n(1)|^{-2}$ under Q. The results below and part (ii) of Proposition 3.9 imply that for these Verblunsky coefficients, the spectral measures μ_α are almost surely singular with respect to the Lebesgue measure.

PROPOSITION 4.3. *With Q probability one*

$$\lim_{n\to\infty} \frac{\log |\varphi_n(1)|^{-2}}{\log n} = -\frac{2}{\beta} = \lim_{n\to\infty} \frac{\log |\psi_n(1)|^{-2}}{\log n}.$$

PROOF. We begin with the proof for $|\varphi_n(1)|^{-2}$. Using equation (4.5) it is enough to determine the asymptotics of the first two summands. Recall that under Q the modified Verblunskies γ_k are independent and have the same law as the original Verblunskies α_k. Therefore

$$\sum_{k=0}^{n-1} \Re\gamma_k$$

is a Q-martingale. The moment assumptions (4.1) and (4.2) show that it satisfies the conditions of Lemma 4.2 for $b_n = \log n$, and therefore is almost surely $o(\log n)$. For the second part, write

$$\sum_{k=0}^{n-1} (\Im\gamma_k)^2 = \sum_{k=0}^{n-1} [(\Im\gamma_k)^2 - \mathbb{E}_Q[(\Im\gamma_k)^2]] + \sum_{k=0}^{n-1} \mathbb{E}_Q[(\Im\gamma_k)^2].$$

By construction the first summation on the right is a martingale, and again by the moment assumptions (1.5) and Lemma 4.2 it is $o(\log n)$ with probability one. For the second summation on the right, we have

$$\mathbb{E}_Q[(\Im\gamma_k)^2] = \frac{1}{2}\mathbb{E}_Q[|\gamma_k|^2] \sim \frac{1}{\beta k}.$$

The first equality is by the rotation invariance of γ_k and the second one is by (4.1). After collecting all signs and the necessary factors of two the proof for $|\varphi_n(1)|^{-2}$ is complete. The result for $|\psi_n(1)|^{-2}$ follows since $\psi_n(1; \boldsymbol{\alpha}) = \varphi_n(1; -\boldsymbol{\alpha})$ and the sequence $-\boldsymbol{\alpha}$ has the same law as $\boldsymbol{\alpha}$ under Q. $\square$

4.5. *Asymptotics under Q_0.*

4.5.1. *First kind polynomials.* Under Q_0 the proofs for $|\varphi_n(1)|^{-2}$ and $|\psi_n(1)|^{-2}$ are very different, in contrast with the computations under Q. We begin with $|\varphi_n(1)|^{-2}$, for which the first summand of (4.5) is decomposed into

$$\sum_{k=0}^{n-1} \Re\gamma_k = \sum_{k=0}^{n-1} (\Re\gamma_k - \mathbb{E}_{Q_0}[\Re\gamma_k]) + \sum_{k=0}^{n-1} \mathbb{E}_{Q_0}[\Re\gamma_k].$$

By construction, the first summation is a martingale under Q_0, and by (4.1), it satisfies the hypotheses of Lemma 4.2 for $b_n = \log n$, since the law of $|\gamma_k|$ is the same under both Q and Q_0 (see Proposition 3.15). Therefore we only need to consider the second summation, for which we have

$$(4.6) \qquad \mathbb{E}_{Q_0}[\Re\gamma_k] = \Re\mathbb{E}_{Q_0}[\gamma_k] = \mathbb{E}_Q[|\gamma_k|^2],$$

the second equality coming from Proposition 3.15. Therefore by the moment assumptions (4.1)

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=0}^{n-1} 2\Re\gamma_k = \frac{4}{\beta},$$

with Q_0-probability one. For the summation involving $(\Im\gamma_k)^2$ the same argument applies, one only needs to use the identity $2(\Im z)^2 = |z|^2 - \Re(z^2)$ and Proposition 3.15 to compute that

$$(4.7) \qquad 2\mathbb{E}_{Q_0}[(\Im\gamma_k)^2] = \mathbb{E}_Q[|\gamma_k|^2] - \mathbb{E}_Q[|\gamma_k|^4] \sim \frac{2}{\beta k}.$$

Therefore

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=0}^{n-1} 2(\Im\gamma_k)^2 = \lim_{n\to\infty} \frac{1}{\log n} \sum_{k=0}^{n-1} E_Q[|\gamma_k|^2] - E_Q[|\gamma_k|^4] = \frac{2}{\beta},$$

Q_0-almost surely.

4.5.2. *Second kind polynomials.* We now study the second kind polynomials ψ_n, and we wish to show that

$$\lim_{n\to\infty} \frac{\log |\psi_n(1)|^{-2}}{\log n} = -\frac{2}{\beta}.$$

Recall that we will always use the assumption (4.3) that the α_n, and thus the γ_n, are uniformly bounded away from 1.

To begin with, for each z define a sequence of matrices $R_n(z)$ by

$$R_n(z) = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_n(z)/\varphi_n(1) & -i\psi_n(z)/\varphi_n(1) \\ \varphi_n^*(z)/\varphi_n^*(1) & i\psi_n^*(z)/\varphi_n^*(1) \end{pmatrix}.$$

The interest in $R_n(z)$ lies in the fact that

$$(4.8) \qquad R_n^*(1)R_n(1) = \begin{pmatrix} 1 & z_n \\ z_n & \left|\dfrac{\psi_n(1)}{\varphi_n(1)}\right|^2 \end{pmatrix}$$

for some quantity z_n that is irrelevant for our purposes. We want to understand the behavior of $|\psi_n(1)|$ under Q_0, and we already know from the previous section that $\log |\varphi_n(1)| \sim$

$-1/\beta \log n$. The Szegő recursion (2.10) and some basic algebra shows that $R_n(z)$ obeys the recursion

$$(4.9) \qquad R_{n+1}(z) = B_n Z R_n(z),$$

where

$$(4.10) \qquad B_n = \begin{pmatrix} (1 - \gamma_n)^{-1} & -\gamma_n(1 - \overline{\gamma_n})^{-1} \\ -\overline{\gamma_n}(1 - \overline{\gamma_n})^{-1} & (1 - \overline{\gamma_n})^{-1} \end{pmatrix}, \qquad Z = \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix}.$$

By iterating (4.9) we can write

$$R_n(1) = B_{n-1} \cdots B_0 U, \qquad U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Since U is a unitary matrix, we have

$$(4.11) \qquad R_n^*(1) R_n(1) = U^*(B_{n-1} \cdots B_0)^*(B_{n-1} \cdots B_0) U = (C_{n-1} \cdots C_0)^* C_{n-1} \cdots C_0,$$

where $C_k = U^* B_k U$. It is straightforward to compute that

$$(4.12) \qquad C_k = \begin{pmatrix} 1 & 2\Im\dfrac{\gamma_k}{1 - \gamma_k} \\ 0 & 1 + 2\Re\dfrac{\gamma_k}{1 - \gamma_k} \end{pmatrix}.$$

REMARK 4.4. The simplification in the shape of the matrices when transforming B_n into C_n can be understood as follows. First, A_n is of the form

$$\begin{pmatrix} x & 1 - x \\ 1 - \overline{x} & \overline{x} \end{pmatrix},$$

and hence corresponds to a Mőbius transformation of the disk. Additionally, the vector $(1\ 1)^t$ is an eigenvector with eigenvalue 1, which means that the point 1 on the disk is left invariant. Now, U corresponds to the Cayley transformation that conformally maps the upper half-plane to the disk, and sends ∞ to 1, and conversely for U^*. The matrix $C_n = U^* A_n U$ therefore represents the image of that Mőbius mapping in the upper half-plane, which therefore fixes ∞ (the image of $(1\ 1)^t$). Hence, it is an affine map, and is therefore written as

$$\begin{pmatrix} 1 & x \\ 0 & y \end{pmatrix},$$

as verified by the above computation.

If we write

$$(4.13) \qquad C_{n-1} \cdots C_0 = \begin{pmatrix} 1 & X_{n-1} \\ 0 & Y_{n-1} \end{pmatrix},$$

then (4.12) shows that X_n and Y_n satisfy the recurrence

$$(4.14) \qquad X_n = X_{n-1} + 2Y_{n-1}\Im\frac{\gamma_n}{1 - \gamma_n}, \qquad Y_n = Y_{n-1}\left(1 + 2\Re\frac{\gamma_n}{1 - \gamma_n}\right).$$

Note that from the expression (2.20) for the Poisson kernel we have the identity

$$(4.15) \qquad Y_{n-1} = \prod_{k=0}^{n-1} P_{\mathbb{D}}(\gamma_k, 1) = |\varphi_n(1)|^{-2}.$$

Finally, (4.8), (4.11), and (4.13) imply that

$$(4.16) \qquad \left| \frac{\psi_n(1)}{\varphi_n(1)} \right|^2 = X_{n-1}^2 + Y_{n-1}^2.$$

Therefore, it is enough to study the sequences $\{X_n\}$ and $\{Y_n\}$ as given by (4.14), under the law Q_0.

For Y_n the analysis is simple enough: by (4.15) and Proposition 4.1 that gives the asymptotics of $|\varphi_n(1)|^{-2}$ under Q_0, we have

$$(4.17) \qquad \lim_{n \to \infty} \frac{\log Y_n}{\log n} = \frac{2}{\beta}, \qquad Q_0 \text{ almost surely.}$$

In particular, for fixed $\varepsilon, \eta > 0$, we have a constant $C > 0$ such that

$$Y_k \le C k^{2/\beta + \eta/2}$$

for all k, with Q_0 probability at least $1 - \varepsilon$.

Now, for the asymptotic behavior of X_n, by (4.14) it is clearly enough to study

$$(4.18) \qquad \sum_{k=1}^{n} Y_{k-1} \Im \frac{\gamma_k}{1 - \gamma_k}$$

under Q_0. In fact we will study the truncated quantity

$$(4.19) \qquad M_n = \sum_{k=1}^{n} (Y_{k-1} \wedge C k^{2/\beta + \eta/2}) \Im \frac{\gamma_k}{1 - \gamma_k}.$$

By the remark above we have that (4.18) and (4.19) are equal with probability at least $1 - \epsilon$. Moreover, the coupling relationship of Proposition 3.15 gives us the identity

$$(4.20) \qquad \mathbb{E}_{Q_0} \left[\Im \frac{\gamma_n}{1 - \gamma_n} \right] = \mathbb{E}_Q \left[\frac{\Im \gamma_n}{1 - |\gamma_n|^2} \right] = 0,$$

with the last equality following from the rotation invariance of γ_n under Q. Hence (4.18) and (4.19) are both mean zero martingales, by the measurability of Y_{k-1} with respect to $\mathcal{F}_{k-1}$. The variance of each term of (4.19) is bounded by

$$(4.21) \qquad \mathbb{E}_{Q_0} \left[(Y_{k-1} \wedge C k^{2/\beta + \eta/2})^2 \left(\Im \frac{\gamma_k}{1 - \gamma_k} \right)^2 \right] \le C^2 k^{4/\beta + \eta} \mathbb{E}_Q \left[\left(\frac{\Im \gamma_n}{1 - |\gamma_n|^2} \right)^2 \right]$$
$$\le C' k^{4/\beta + \eta - 1},$$

where we use (4.3). Now, an appeal to Lemma 4.2 with $b_k = k^{2/\beta + \eta}$ allows us to conclude that

$$\frac{M_n}{n^{2/\beta + \eta}} \to 0, \qquad Q_0\text{-a.s.}$$

Since M_n is equal to (4.18) with probability at least $1 - \epsilon$, and since the recursion (4.14) tells us that (4.18) has the same asymptotics as X_n, we conclude that

$$\lim_{n \to \infty} \frac{X_n}{n^{2/\beta + \eta}} = 0$$

with Q_0 probability at least $1 - \epsilon$. Since this holds for arbitrary ϵ and η, we finish by using (4.16) and (4.17) to conclude that

$$\lim_{n \to \infty} \frac{1}{\log n} \log \left| \frac{\psi_n(1)}{\varphi_n(1)} \right|^2 = \frac{4}{\beta}, \qquad Q_0\text{-a.s.}$$

Since we already proved in Section 4.5.1 that

$$\lim_{n\to\infty} \frac{\log|\varphi_n(1)|^{-2}}{\log n} = \frac{2}{\beta}, \quad Q_0\text{-a.s.},$$

the proof of Proposition 4.1 follows.

5. Large deviations for the norm.   Jitomirskaya–Last dimension theory requires strong control on the norm

$$\|\varphi.(1)\|_n^2 = \sum_{k=0}^{n} |\varphi_k(1)|^2.$$

In the previous section we obtained this control by finding precise asymptotics for $\log|\varphi_n(1)|^2$ as $n \to \infty$, using that it is a sum of independent variables (under both Q and Q_0). From this we derived that

$$\lim_{n\to\infty} \frac{\log\|\varphi.(1)\|_n^2}{\log n} = 1 + \frac{2}{\beta}, \, Q\text{-a.s.}, \qquad \lim_{n\to\infty} \frac{\log\|\varphi.(1)\|_n^2}{\log n} = 1 - \frac{2}{\beta}, \, Q_0\text{-a.s.}$$

In this section we derive a large deviations principle (LDP) for the quantity

$$\Upsilon_n := \frac{\log\|\varphi.(1)\|_n^2}{\log n},$$

that is, the asymptotic probability that Υ_n takes on an atypical value as $n \to \infty$. Such probabilities are useful for estimating the probability that the spectral measure of an interval, $\mu_{\boldsymbol{\alpha}}(\theta - \epsilon, \theta + \epsilon)$, decays atypically as $\epsilon \to 0$, for θ chosen according to $\mu_{\boldsymbol{\alpha}}$ or to the Lebesgue measure. The large deviations analysis is somewhat delicate because the scale of the terms $|\varphi_k(1)|^2$ changes as k increases. Moreover we need to get control on the large deviations behavior of the entire process $k \mapsto |\varphi_k(1)|^2$ in order to get control on the behavior of the norm $\|\varphi.(1)\|_n^2$. To this end we analyze the large deviations of the sequence of processes

(5.1)
$$Z_n(t) := \frac{\log|\varphi_{k_n(t)}(1)|^{-2}}{\log n}, \quad t \in [0, 1],$$

as random elements of the Skorohod space $\mathbb{D}([0, 1])$, for an appropriately chosen time scale k_n. We obtain a functional LDP for the sequence (Z_n) and convert it into an LDP for (Υ_n), using the fact that $\|\varphi.(1)\|_n^2$ is a functional of $(|\varphi_k(1)|^2)$, applying the *contraction principle*. Process level LDPs are often easier to obtain, and that is also true in this case because $\log|\varphi_k(1)|^{-2}$ is the sum of independent random variables. The process level LDP for (Z_n) is also helpful for understanding how the Verblunsky coefficients behave when they produce an atypical value of (Υ_n).

5.1. *Process level LDP for sums of independent variables.*   The standard process level result for sums of *i.i.d.* variables is Mogulskii's theorem; see [14], Chapter 5. Here is a quick summary: assume the i.i.d. random variables $X_1, X_2, \ldots$, taking values in $\mathbb{R}$, satisfy that

$$\Lambda(\lambda) = \log \mathbb{E}[e^{\lambda X_i}]$$

is finite for all $\lambda \in \mathbb{R}$. Define the process $S_n : [0, 1] \to \mathbb{R}$ by

$$S_n(t) = \frac{1}{n} \sum_{i=1}^{nt} X_i,$$

which we consider as a random element of the Skorohod space $\mathbb{D}([0, 1])$. Then Mogulskii's theorem states that the sequence of probability laws induced by S_n satisfies an LDP with speed n and rate function

$$(5.2) \qquad I(g) = \begin{cases} \int_0^1 \Lambda^*(g(t))\,dt, & \text{if } g \text{ is absolutely continuous and } g(0) = 0, \\ +\infty, & \text{otherwise,} \end{cases}$$

where Λ^* is the Fenchel–Legendre transform of Λ defined by

$$\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}} \{\lambda x - \Lambda(\lambda)\}.$$

Recall that the precise statement of the LDP is that for every Borel set $A \subset \mathbb{D}([0, 1])$, we have

$$-\inf_{g \in A^\circ} I(g) \le \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}(S_n \in A) \le \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(S_n \in A) \le -\inf_{g \in \overline{A}} I(g)$$

where A° and $\overline{A}$ are the interior and closure of A, respectively, in the Skorohod topology.

The process level large deviations principle for $\log |\varphi_n(1)|^{-2}$ does not fall exactly into this framework, since by equation (2.23) we have

$$\log|\varphi_n(1)|^{-2} = \sum_{k=0}^{n-1} \log P_{\mathbb{D}}(\gamma_k, 1).$$

The γ_k are still independent (under both Q and Q_0) but the distribution of the γ_k is changing and concentrating around zero as $k \to \infty$. Thus we need a version of Mogulskii's theorem in which the random variables X_i are independent but not necessarily identically distributed. This should be a standard result but we were unable to find it in the literature, so we spend the rest of this section deriving it. Most of the proof is a modification of the one for Mogulskii's theorem in [18], Chapter 4, with some modifications from [14], Chapter 5. The statement is the following.

THEOREM 5.1. *Let (X_k) be a sequence of independent random variables taking values in $\mathbb{R}$ such that*

$$\Lambda_k(\lambda) := \log \mathbb{E}[e^{\lambda X_k}]$$

is finite in an open interval around zero. Moreover, assume there is a sequence of positive numbers (c_k) such that

$$\Lambda(\lambda) := \lim_{k \to \infty} \frac{1}{c_k} \Lambda_k(\lambda)$$

exists and is finite for every $\lambda \in \mathbb{R}$. Further assume that (c_k) and Λ satisfy:

(a) *$\log n = O(K_n)$ and $\max_{k \le n} c_k = o(K_n)$, where $K_n := c_1 + \cdots + c_n$;*
(b) *the function Λ is differentiable and steep, the latter meaning that*

$$\lim_{\lambda \to \pm\infty} \Lambda'(\lambda) = \infty.$$

Now, for $n \ge 1$ and $1 \le k \le n$, define a mesh $t_{k,n}$ in $[0, 1]$ by

$$t_{k,n} = (c_1 + \cdots + c_k)/K_n,$$

and $t_{0,n} = 0$. Then define a time scale $k_n : [0, 1] \to \{1, \ldots, n\}$ by $k_n(t) = k$ for $t \in [t_{k-1,n}, t_{k,n})$, and consider the partial sum process

$$(5.3) \qquad Z_n(t) = \frac{1}{K_n} \sum_{i=1}^{k_n(t)} X_i, \quad t \in [0, 1].$$

Then the processes Z_n satisfy a LDP with speed K_n and rate function (5.2), *in the topology of uniform convergence.*

Note that the process $Z_n(t)$ of (5.1) is just a special case of (5.3) with $X_k = \log P_{\mathbb{D}}(\gamma_k, 1)$ and c_k still to be determined; we will later see that the appropriate choice is $c_k = 1/k$. For many choices of Verblunsky coefficients and k small it will often be true that

$$\mathbb{E}[P_{\mathbb{D}}(\gamma_k, 1)^{\lambda}] = \infty,$$

at least for λ sufficiently large, which is why we do not assume that the Λ_k are finite for all λ. However as long as the Λ_k eventually become finite at each λ and converge to some limiting value there are no difficulties. We will verify that the conditions of Theorem 5.1 hold for a broad class of Verblunsky coefficients in the next section. The steepness condition on Λ ensures that its Fenchel–Legendre transform Λ^* is a good convex rate function (see [14], Theorem 2.3.6), which is implicitly used in the proofs to come. Also note that the topology of the theorem is uniform convergence, even though the processes Z_n have jumps at the times $t_{k,n}$ and are therefore in the Skorohod space $\mathbb{D}([0, 1])$. Thus one of the conclusions of the theorem is that the jumps can be ignored, or the process can be made continuous by linear interpolation without any modification to the LDP. We will use the latter fact later to obtain the LDP for Υ_n.

We break the proof of Theorem 5.1 into three lemmas. In the first, Lemma 5.3, we show that the processes Z_n are exponentially tight in the Skorohod space $\mathbb{D}([0, 1])$. In the second lemma we show that the Z_n are C-exponentially tight. This allows us to prove the LDP in the topology of uniform convergence rather than just the Skorohod topology. The final lemma proves an LDP for finite-dimensional distributions of the process, and then combined with exponential tightness, this allows us to infer the process-level LDP. The latter is essentially a version of the Dawson–Gärtner theorem, although we use the approach of [18], Chapter 4.7. In these lemmas we need the following straightforward extension of the Stolz–Cesàro theorem, whose proof is an easy exercise.

LEMMA 5.2. *Assume that (c_k) is a sequence of positive numbers with $K_n = c_1 + \cdots + c_n \to \infty$. Assume that (d_k) is another sequence of numbers with $d_n/c_n \to \ell \in \mathbb{R}$ as $n \to \infty$, and that there are integer sequences $(r_n^{\pm})$ such that*

$$\lim_{n \to \infty} \frac{1}{K_n} \sum_{i=r_n^-}^{r_n^+} c_i = u.$$

Then

$$\lim_{n \to \infty} \frac{1}{K_n} \sum_{i=r_n^-}^{r_n^+} d_i = u\ell.$$

LEMMA 5.3. *The laws of Z_n are exponentially tight in $\mathbb{D}([0, 1])$, meaning that for every $M < \infty$, there exists a compact set $E_M \subset \mathbb{D}([0, 1])$ such that*

$$\limsup_{n \to \infty} \frac{1}{K_n} \log \mathbb{P}(Z_n \in E_M) < -M.$$

PROOF. The proof is based on [18], Theorem 4.1. We first show that, for each fixed $t \in [0, 1]$, the sequence $(Z_n(t))$ is exponentially tight in $\mathbb{R}$. Let $k_n = k_n(t)$, so that

$$Z_n(t) = \frac{1}{K_n} \sum_{i=1}^{k_n} X_i.$$

Then for $L > 0$, the exponential Chebyshev inequality implies that

$$\frac{1}{K_n} \log \mathbb{P}(Z_n(t) > L) \leq -L + \frac{1}{K_n} \sum_{i=1}^{k_n} \Lambda_i(1).$$

By definition of K_n and the assumption that $\Lambda_i(1)/c_i \to \Lambda(1)$ as $i \to \infty$, it follows from the Stolz–Cèsaro type theorem in Lemma 5.2 that the right hand side tends to $t\Lambda(1)$ as $n \to \infty$. The same argument produces a similar L dependent bound for $\mathbb{P}(Z_n(t) < -L)$, and then taking $L \to \infty$ shows the exponential tightness.

Now we prove that the Z_n are tight in the Skorohod space $\mathbb{D}([0, 1])$. As in [18], Theorem 4.1, this is done by using the pointwise exponential tightness above and proving that

$$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{1}{K_n} \log \mathbb{P}(\omega'(Z_n, \delta) > \epsilon) = -\infty,$$

where $\omega' : \mathbb{D}([0, 1]) \times \mathbb{R} \to [0, \infty]$ is the modulus-of-continuity type object defined by

$$\omega'(f, \delta) = \inf_{(s_j)} \max_j \sup_{s_{j-1} \leq s, t \leq s_j} |f(s) - f(t)|,$$

and the infimum is over all partitions of $[0, 1]$ with mesh size (minimum length of an interval) greater than δ. We will only need to consider the partition with all intervals the same size and of length δ, except for potentially the last one having longer length in $[\delta, 2\delta)$. For $s < t$ we have by definition of Z_n that

$$Z_n(t) - Z_n(s) = \frac{1}{K_n} \sum_{i=k_n(s)+1}^{k_n(t)} X_i.$$

Therefore, by using the equisized partition mentioned above we have

$$\omega'(Z_n, \delta) \leq \max_{0 \leq j < 1/\delta} \frac{1}{K_n} \left| \sum_{i=k_n(j\delta)+1}^{k_n((j+1)\delta)} X_i \right|.$$

We then apply the standard union bound and replace the summation with the maximum of the summands to obtain

$$\mathbb{P}(\omega'(Z_n, \delta) > \epsilon) \leq \delta^{-1} \max_{0 \leq j < 1/\delta} \mathbb{P}\left(\frac{1}{K_n} \left| \sum_{i=k_n(j\delta)+1}^{k_n((j+1)\delta)} X_i \right| > \epsilon \right).$$

Each of the probabilities above can be bounded by the exponential Chebyshev inequality to obtain

$$\mathbb{P}(\omega'(Z_n, \delta) > \epsilon) \leq \delta^{-1} e^{-\lambda K_n \epsilon} \max_{0 \leq j < 1/\delta} \max_{\pm \lambda} \prod_{i=k_n(j\delta)+1}^{k_n((j+1)\delta)} \mathbb{E} e^{\pm \lambda X_i},$$

where the inner maximum means the larger over the indicated terms with λ or $-\lambda$. Therefore by taking logarithms we have

$$(5.4) \qquad \frac{1}{K_n} \log \mathbb{P}(\omega'(Z_n, \delta) > \epsilon) \leq -\frac{\log \delta}{K_n} - \lambda \epsilon + \max_{0 \leq j < 1/\delta} \max_{\pm \lambda} \frac{1}{K_n} \sum_{i=k_n(j\delta)+1}^{k_n((j+1)\delta)} \Lambda_i(\pm \lambda).$$

However, by definition of K_n and the functions k_n, we have

$$\lim_{n \to \infty} \frac{1}{K_n} \sum_{i=k_n(j\delta)+1}^{k_n((j+1)\delta)} c_i = (j+1)\delta - j\delta = \delta,$$

and since $\Lambda_i(\pm\lambda)/c_i \to \Lambda(\pm\lambda)$ as $i \to \infty$, the Stolz–Cèsaro theorem of Lemma 5.2 implies that

$$\lim_{n\to\infty} \frac{1}{K_n} \sum_{i=k_n(j\delta)+1}^{k_n((j+1)\delta)} \Lambda_i(\pm\lambda) = \delta\Lambda(\pm\lambda).$$

Therefore by taking the limsup of (5.4) we obtain

$$\limsup_{n\to\infty} \frac{1}{K_n} \sum_{i=k_n(j\delta)+1}^{k_n((j+1)\delta)} \Lambda_i(\pm\lambda) \le -\lambda\epsilon + \delta \max_{\pm\lambda} \Lambda(\pm\lambda) = -\delta \min_{\pm\lambda}\{\lambda\epsilon/\delta - \Lambda(\pm\lambda)\}.$$

This is true for all $\lambda > 0$, therefore we can optimize the right hand side over λ to obtain

$$(5.5) \qquad \limsup_{n\to\infty} \frac{1}{K_n} \log \mathbb{P}(\omega'(Z_n, \delta) > \epsilon) \le -\delta \min_{\pm\lambda} \Lambda^*(\pm\epsilon/\delta).$$

Then the steepness condition $\Lambda'(\lambda) \to \infty$ as $\lambda \to \pm\infty$ implies that $\delta\Lambda^*(\pm\epsilon/\delta) \to \infty$ as $\delta \to 0$, so the above exactly shows that the Z_n are exponentially tight in the Skorohod space $\mathbb{D}([0, 1])$. $\square$

LEMMA 5.4. *The processes Z_n are C-exponentially tight in $D([0, 1])$, meaning that for every $\delta > 0$*

$$\limsup_{n\to\infty} \frac{1}{K_n} \log \mathbb{P}\Big(\sup_{t\in[0,1]} |Z_n(t) - Z_n(t-)| > \delta \Big) = -\infty.$$

PROOF.   By definition of Z_n the only jumps are at the times $t_{k,n}$, therefore

$$\sup_{t\in[0,1]} |Z_n(t) - Z_n(t-)| = \frac{1}{K_n} \max_{1\le k\le n} |X_k|.$$

Then by the standard union bound

$$\mathbb{P}\Big(\sup_{t\in[0,1]} |Z_n(t) - Z_n(t-)| > \delta \Big) \le \sum_{k=1}^{n} \mathbb{P}(|X_k| > \delta K_n).$$

Now, for each fixed $\lambda > 0$, there exists an integer $M = M(\lambda)$ such that $\mathbb{E}[e^{\lambda|X_k|}] < \infty$ for all $k \ge M(\lambda)$, by the assumption that $\Lambda_k(\pm\lambda)/c_k$ converges to a finite quantity for each λ. Therefore on all terms with $k \ge M$ in the last summation we can apply the exponential Chebyshev inequality to obtain

$$\sum_{k=1}^{n} \mathbb{P}(|X_k| > \delta K_n) \le M + \sum_{k=M}^{n} e^{-\lambda\delta K_n} \mathbb{E}[e^{\lambda|X_k|}].$$

Taking logarithms of both sides yields (also using subadditivity of the logarithm)

$$\frac{1}{K_n} \log \mathbb{P}\Big(\sup_{t\in[0,1]} |Z_n(t) - Z_n(t-)| > \delta \Big) \le \frac{1}{K_n} \log M - \delta\lambda + \frac{1}{K_n} \log \sum_{k=M}^{n} \mathbb{E}[e^{\lambda|X_k|}].$$

By the assumption that $c_k^{-1}\Lambda_k(\pm\lambda)$ converges as $k \to \infty$ we have that there exists a constant $A = A(\lambda) > 0$ such that $\log \mathbb{E}[e^{\pm\lambda X_k}] \le Ac_k$ for all $k \ge M$, and therefore

$$\frac{1}{K_n} \log \mathbb{P}\Big(\sup_{t\in[0,1]} |Z_n(t) - Z_n(t-)| > \delta \Big) \le \frac{1}{K_n} \log M - \delta\lambda + \frac{2}{K_n} \log \sum_{k=M}^{n} e^{Ac_k}$$

$$\le \frac{1}{K_n} \log M - \delta\lambda + \frac{2}{K_n}\Big(A \max_{1\le k\le n} c_k + \log n\Big).$$

Now, recalling that $K_n = c_1 + \cdots + c_n$ and that we assumed that $\log n = O(K_n)$ and $\max_{k \le n} c_k = o(K_n)$, for the last term in the expression above there is a universal constant $C > 0$ such that

$$\limsup_{n \to \infty} \frac{1}{K_n} \log \mathbb{P}\Big(\sup_{t \in [0,1]} |Z_n(t) - Z_n(t-)| > \delta \Big) \le -\delta\lambda + C.$$

The latter holds for all $\lambda > 0$, so taking $\lambda \to \infty$ completes the proof. $\square$

Note that to handle the case that $\Lambda_k(\lambda) = \infty$ for k small required only mild modifications. In the remaining two lemmas no extra modifications beyond those used above are necessary to handle this complication, so we provide the proofs under the assumption $\Lambda_k(\lambda) < \infty$ for all k and λ and let the reader fill in the details. For the full details of how to use Lemma 5.4 to strengthen the topology of the LDP for Z_n to that of uniform convergence; see [18], Chapter 4.4.

LEMMA 5.5. *Let $s_1 < s_2 < \cdots s_m$ be a partition in $[0, 1]$. Then the process $(Z_n(s_1), \ldots, Z_n(s_m))$ satisfies a large deviations principle in $\mathbb{R}^m$ with speed K_n and rate function*

$$I^Z_{s_1,\ldots,s_m}(\mathbf{x}) = \sum_{j=1}^{m} (s_j - s_{j-1})\Lambda^*\left(\frac{\mathbf{x}_j - \mathbf{x}_{j-1}}{s_j - s_{j-1}}\right).$$

Consequently, Z_n satisfies a LDP with speed K_n and rate function (5.2), *in the topology of uniform convergence.*

PROOF OF LEMMA 5.5. The proof is similar to [14], Lemma 5.1.6, which proves the same statement for the standard situation of i.i.d. random variables. In our situation all that needs to be modified is the proof of the finite-dimensional LDP which we now give. Let Y_n be the random element of $\mathbb{R}^m$ given by

$$Y_n = (Z_n(s_1), Z_n(s_2) - Z_n(s_1), \ldots, Z_n(s_m) - Z_n(s_{m-1})).$$

Further, for n large let $0 = t_{k_0,n} < t_{k_1,n} < \cdots < t_{k_m,n} \le 1$ be the ordered sequence of times that are closest to the s_j in the mesh, in the sense that $s_j \in [t_{k_j,n}, t_{k_j+1,n})$ for all $1 \le j \le m$. Since the mesh size goes to zero as $n \to \infty$ (by $\max_{k \le n} c_k = o(K_n)$), this is always possible (with the $t_{k_i,n}$ distinct) for n large enough. Also note that the terms k_j depend on n but we suppress the dependence in the notation. Now writing $Y_n^j = Z_n(s_j) - Z_n(s_{j-1})$ for the coordinates of Y_n we have by definition of Z_n that

$$Y_n^j = Z_n(s_j) - Z_n(s_{j-1}) = Z_n(t_{k_j,n}) - Z_n(t_{k_{j-1},n+1}) = \frac{1}{K_n} \sum_{i=k_{j-1}+1}^{k_j} X_i.$$

Therefore, for $(\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$, the exponential moment-generating function of the variable $Y_n = (Y_n^1, \ldots, Y_n^m)$ satisfies

$$(5.6) \qquad \log \mathbb{E} \exp\left(K_n \sum_{j=1}^{m} \lambda_j Y_n^j \right) = \sum_{j=1}^{m} \sum_{i=k_{j-1}+1}^{k_j} \Lambda_i(\lambda_j).$$

But by definition of the series $t_{k,n}$ we have that

$$t_{k_j,n} - t_{k_{j-1},n} = \frac{1}{K_n} \sum_{i=k_{j-1}+1}^{k_j} c_i,$$

and since the $t_{k_j,n}$ are the closest points to s_j in the mesh, we have that the above converges to $s_j - s_{j-1}$ as $n \to \infty$. Furthermore, by the definition of Λ as a limit, we have that $\Lambda_i(\lambda_j)/c_i \to \Lambda(\lambda_j)$ as $i \to \infty$, so by combining these last two facts and Lemma 5.2, we have

$$\lim_{n \to \infty} \frac{1}{K_n} \sum_{i=k_{j-1}+1}^{k_j} \Lambda_i(\lambda_j) = (s_j - s_{j-1})\Lambda(\lambda_j).$$

Combining this with (5.6) we obtain that

$$\lim_{n \to \infty} \frac{1}{K_n} \log \mathbb{E} \exp\left(K_n \sum_{j=1}^{m} \lambda_j Y_n^j \right) = \sum_{j=1}^{m} (s_j - s_{j-1})\Lambda(\lambda_j).$$

By the assumptions on Λ the right hand side is, as a function of $(\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$, differentiable, lower semi-continuous, and steep (in any direction going to infinity in $\mathbb{R}^m$). Therefore, by the Gärtner–Ellis theorem [14], Theorem 2.3.6, the processes Y_n satisfy an LDP as random elements of $\mathbb{R}^m$, with speed K_n and rate function

$$I^Y_{s_1,\ldots,s_m}(\mathbf{x}) = \sum_{j=1}^{m} (s_j - s_{j-1})\Lambda^*\left(\frac{\mathbf{x}_j}{s_j - s_{j-1}} \right)$$

for $\mathbf{x} \in \mathbb{R}^m$. Finally, since the mapping $Y_n = (Z_n(s_1), Z_n(s_2) - Z_n(s_1), \ldots, Z_n(s_m) - Z_n(s_{m-1})) \mapsto (Z_n(s_1), \ldots, Z_n(s_m))$ is clearly continuous and bijective, the LDP for Y_n translates into an LDP for Z_n with speed K_n and rate function

$$I^Z_{s_1,\ldots,s_m}(\mathbf{x}) = \sum_{j=1}^{m} (s_j - s_{j-1})\Lambda^*\left(\frac{\mathbf{x}_j - \mathbf{x}_{j-1}}{s_j - s_{j-1}} \right),$$

by means of the contraction principle [14], Theorem 4.2.1. This completes the LDP for the processes $(Z_n(s_1), \ldots, Z_n(s_m))$. To extend to the process-level LDP, use the C-exponential tightness of Lemma 5.3 and [18], Theorem 4.30, to conclude that Z_n satisfies the LDP with good rate function

$$I(g) = \sup_{\{s_i\}} I_{s_1,\ldots,s_m}\big(g(s_1), \ldots, g(s_m) \big),$$

where the supremum is over all partitions in $[0, 1]$ (for any value of m). To conclude that this supremum is equal to the rate function $I(g)$ of (5.2) is a straightforward calculus exercise; see [14], Corollary 5.1.10, for full details. $\quad\square$

5.2. *LDP for Z_n.* Now we apply the results of the last section to the orthogonal polynomials. The next lemma verifies the assumptions of Theorem 5.1 for the specific partial sum process of (5.1), and identifies the corresponding rate functions.

LEMMA 5.6. *Assume that under the measure Q, the Verblunsky coefficients α_k are independent and rotation invariant with the usual second moment assumption (1.5). Further assume that for some $\epsilon > 0$ and for all $\kappa > 0$, the radial parts satisfy*

$$\mathbb{E}\big[|\alpha_k|^3\big] = O\big(k^{-1-\epsilon}\big), \qquad \limsup_{k \to \infty} \mathbb{E}\big[(1 - |\alpha_k|)^{-\kappa}\big] < \infty.$$

Then $X_k = \log P_{\mathbb{D}}(\gamma_k, 1)$ satisfies the hypotheses of Theorem 5.1 with $c_k = 1/k$ and

$$\Lambda^*(x) = \frac{\beta}{8}\left(x + \frac{2}{\beta} \right)^2.$$

Furthermore, under Q_0, the same X_k satisfy the hypotheses also with $c_k = 1/k$ and rate function

$$\Lambda^*(x) - x = \frac{\beta}{8}\left(x - \frac{2}{\beta}\right)^2.$$

PROOF. First note that rotation invariance of the α_k implies that the modified Verblunskies γ_k are independent (Lemma 3.2), but that is all we use the rotation invariance for. After that we only need to check that for each $\lambda \in \mathbb{R}$ the limit

$$\lim_{k\to\infty} \frac{1}{c_k} \log \mathbb{E}\big[\exp\{\lambda \log P_{\mathbb{D}}(\gamma_k, 1)\}\big] = \lim_{k\to\infty} k \log \mathbb{E}\big[P_{\mathbb{D}}(\gamma_k, 1)^\lambda\big] =: \Lambda(\lambda)$$

exists and is finite, and that the resulting Λ is differentiable and steep. First work on the event that $|\alpha_k| \leq 1/2$, where the truncation allows us to use the expansion (4.4) to write

$$P_{\mathbb{D}}(\gamma_k, 1)^\lambda = 1 + 2\lambda \Re\gamma_k - 2\lambda(\Im\gamma_k)^2 + 2\lambda^2(\Re\gamma_k)^2 + O(|\gamma_k|^3),$$

and the constants in the O term are uniform in α_k. We already saw in Section 4.2 that

$$\mathbb{E}\big[|\gamma_k|^2 \mathbf{1}_{\{|\gamma_k| \leq 1/2\}}\big] = \frac{2}{\beta k} + o(k^{-1}),$$

so that we may take an expectation of the above asymptotic expression to obtain

$$(5.7) \qquad \mathbb{E}\big[P_{\mathbb{D}}(\gamma_k, 1)^\lambda \mathbf{1}_{\{|\gamma_k| \leq 1/2\}}\big] = 1 + \lambda(\lambda - 1)\frac{2}{\beta k} + o(k^{-1-\epsilon}).$$

Now for the case $|\alpha_k| \geq 1/2$ begin with the estimate

$$P_{\mathbb{D}}(z, 1) = 2\Re\frac{1+z}{1-z} \leq \frac{1+|z|}{1-|z|} \leq \frac{2}{1-|z|}.$$

Let $q \in (1, 1+\epsilon)$ and p satisfy $1/p + 1/q = 1$. Then from the bound above and a combination of Hölder's and Markov's inequality,

$$\begin{aligned}
\mathbb{E}\big[P_{\mathbb{D}}(\gamma_k, 1)^\lambda \mathbf{1}_{\{|\gamma_k| \geq 1/2\}}\big] &\leq C\mathbb{E}\big[(1-|\gamma_k|)^{-\lambda}\mathbf{1}_{\{|\gamma_k| \geq 1/2\}}\big] \\
&\leq C\mathbb{E}\big[(1-|\gamma_k|)^{-\lambda p}\big]^{1/p}\mathbb{P}(|\gamma_k| > 1/2)^{1/q} \\
&\leq C\mathbb{E}\big[(1-|\gamma_k|)^{-\lambda p}\big]^{1/p}\mathbb{E}[|\gamma_k|^3]^{1/q} \\
&\leq C\mathbb{E}\big[(1-|\gamma_k|)^{-\lambda p}\big]^{1/p}O(k^{-(1+\epsilon)/q}),
\end{aligned}$$

the last inequality following by the assumption on $\mathbb{E}[|\gamma_k|^3]$, where C is a constant independent of k. Therefore by the negative moment assumption on $(1 - |\gamma_k|)$ and the fact that $(1 + \epsilon)/q > 1$ we have

$$(5.8) \qquad \limsup_{k\to\infty} k\mathbb{E}\big[P_{\mathbb{D}}(\gamma_k, 1)^\lambda \mathbf{1}_{\{|\gamma_k| \geq 1/2\}}\big] = 0.$$

Combining (5.8) with (5.7) we therefore obtain

$$\Lambda(\lambda) = \lim_{k\to\infty} k \log \mathbb{E}\big[P_{\mathbb{D}}(\gamma_k, 1)^\lambda\big] = \frac{2}{\beta}\lambda(\lambda - 1).$$

This Λ is clearly differentiable and steep, and it is straightforward to verify that its Legendre–Fenchel transform is Λ^* as given. Finally, to prove the statement under Q_0, recall that Theorem 3.14 implies that the γ_k are also independent under Q_0, and the fact that the Radon–Nikodym derivative between Q and Q_0 (for γ_k) is $P_{\mathbb{D}}(\gamma_k, 1)$ gives that

$$\lim_{k\to\infty} k\mathbb{E}_{Q_0}\big[P_{\mathbb{D}}(\gamma_k, 1)^\lambda\big] = \lim_{k\to\infty} k\mathbb{E}_Q\big[P_{\mathbb{D}}(\gamma_k, 1)^{\lambda+1}\big] = \Lambda(\lambda + 1).$$

Therefore Theorem 5.1 also applies under Q_0 with rate function as given. $\quad\square$

5.3. *LDP for Υ_n.* To obtain the LDP for Υ_n we write it as a function of the process Z_n and then apply the contraction principle. Recall that $Z_n(t)$ is defined by (5.1), which can be rewritten as

$$|\varphi_{k_n(t)}|^{-2} = n^{Z_n(t)}, \quad t \in [0, 1].$$

Now recall the definition of the norm $\|\varphi.\|_n^2$ as

$$\|\varphi.(1)\|_n^2 = \sum_{k=0}^{n} |\varphi_k(1)|^2.$$

In the remainder of this section it will be convenient to start the sum at $k = 1$ instead. Clearly, this will have no effect on the large deviations principle. Now using that $k_n(t) = k$ on each interval $t \in [t_{k-1,n}, t_{k,n})$ and $t_{k,n} - t_{k-1,n} = c_k = 1/k$ in our case, we can rewrite the norm as

$$\|\varphi.(1)\|_n^2 = \sum_{k=1}^{n} \frac{1}{c_k} \int_{t_{k-1,n}}^{t_{k,n}} n^{-Z_n(s)} \, ds = K_n \sum_{k=1}^{n} k \int_{t_{k-1,n}}^{t_{k,n}} n^{-Z_n(s)} \, ds = K_n \int_0^1 k_n(s) n^{-Z_n(s)} \, ds.$$

Now by definition of $k_n(s)$ we have

$$k_n(s) = \min\left\{ k \geq 1 : \frac{1}{K_n} \sum_{i=1}^{k} \frac{1}{i} \geq s \right\},$$

from which it follows by the asymptotics of the harmonic series that $k_n(s) = C(s) n^{s+O(1/n)}$, where $C(s)$ is a nonrandom, positive, continuous function on $[0, 1]$ that is strictly bounded away from zero. Therefore we have

$$\|\varphi.(1)\|_n^2 \sim K_n \int_0^1 n^{s-Z_n(s)} ds, \quad n \to \infty,$$

so it is enough to prove the LDP for

$$\frac{1}{\log n} \log \int_0^1 n^{s-Z_n(s)} ds,$$

the latter using that $K_n \sim \log n$. By Laplace's principle we expect that

$$\frac{1}{\log n} \log \int_0^1 n^{s-Z_n(s)} ds \sim \max_{s \in [0,1]} \{s - Z_n(s)\},$$

as $n \to \infty$. In Lemma 5.8 we will show that these two sequences are close enough to be exponentially equivalent, meaning that if one satisfies an LDP then the other satisfies the same LDP. Therefore to complete the proof of Theorem 1.6 it will be enough to prove an LDP for the sequence

$$\max_{s \in [0,1]} \{s - Z_n(s)\},$$

which we do via Theorem 5.1 and the contraction principle.

LEMMA 5.7. *The sequence $\max_{0 \leq s \leq 1} \{s - Z_n(s)\}$ satisfy an LDP with speed K_n and rate function $\Lambda^*(1 - x)$.*

PROOF. Since the mapping $g \to \max_{0 \leq s \leq 1} \{s - g(s)\}$ is continuous from $C_0([0, 1])$ to $\mathbb{R}$, it follows by Theorem 5.1 and the contraction principle that $\max_{0 \leq s \leq 1} \{s - Z_n(s)\}$ satisfies the LDP with speed K_n and rate function

$$J(x) = \inf\left\{ \int_0^1 \Lambda^*(\phi'(t)) \, dt : \max_{0 \leq s \leq 1} \{s - \phi(s)\} = x \right\}.$$

As $\phi(0) = 0$, there is no solution for $x < 0$, and thus $J(x) = \infty$. We therefore now assume that $x \geq 0$. By taking $\phi(s) = s(1 - x)$ we clearly get that $J(x) \leq \Lambda^*(1 - x)$. However, this ϕ is also the minimizer. To see this, let us assume that $s - \phi(s)$ has its maximum at $s_0 \in (0, 1)$ (with trivial modifications if $s_0 \in \{0, 1\}$), so in particular $s_0 - \phi(s_0) = x$. The convexity of Λ^* and Jensen's inequality then give

$$\int_0^1 \Lambda^*(\phi'(s)) \, ds = \int_0^{s_0} \Lambda^*(\phi'(s)) \, ds + \int_{s_0}^1 \Lambda^*(\phi'(s)) \, ds$$

$$\geq s_0 \Lambda^*\left(\frac{1}{s_0} \int_0^{s_0} \phi'(s) \, ds\right) + (1 - s_0)\Lambda^*\left(\frac{1}{1 - s_0} \int_{s_0}^1 \phi'(s) \, ds\right)$$

$$= s_0 \Lambda^*\left(1 - \frac{x}{s_0}\right) + (1 - s_0)\Lambda^*\left(\frac{1}{1 - s_0}(\phi(1) - \phi(s_0))\right).$$

Now note that

$$\frac{1}{1 - s_0}(\phi(1) - \phi(s_0)) \geq \frac{1}{1 - s_0}(1 - x - (s_0 - x)) = 1,$$

and that Λ^* (as well as $\Lambda^*(x) - x$, for $\beta > 2$, when we work under Q_0) is increasing on $[1, \infty)$. From this fact, the previous computations, and the convexity of Λ^*, we therefore deduce that

$$\int_0^1 \Lambda^*(\phi'(s)) \, ds \geq s_0 \Lambda^*\left(1 - \frac{x}{s_0}\right) + (1 - s_0)\Lambda^*(1) \geq \Lambda^*\left(s_0\left(1 - \frac{x}{s_0}\right) + 1 - s_0\right) = \Lambda^*(1 - x),$$

which shows that $\phi(s) = s(1 - x)$ is indeed the minimizer of the functional, and thus

$$J(x) = \Lambda^*(1 - x),$$

as we wanted to show. □

LEMMA 5.8. *The sequences* $\log \int_0^1 n^{s - Z_n(s)} \, ds / \log n$ *and* $\max_{s \in [0,1]}\{s - Z_n(s)\}$ *are exponentially equivalent, meaning that for every* $\delta > 0$,

$$\limsup_{n \to \infty} \frac{1}{K_n} \log \mathbb{P}\left(\left|\frac{1}{\log n} \log \int_0^1 n^{s - Z_n(s)} ds - \max_{s \in [0,1]}\{s - Z_n(s)\}\right| > \delta\right) = -\infty.$$

PROOF. Throughout we will write $Y_n(s) = s - Z_n(s)$. We will use the continuity properties of Y_n to prove the equivalence, along the following easily derived bounds for Laplace's principle: if f is a bounded, measurable function on $[0, 1]$ and if

$$\sup_{|s-t|<\delta} |f(s) - f(t)| < \epsilon$$

for some particular $\epsilon > 0$ and $\delta \in (0, 1/2)$, then

(5.9) $$0 \leq \operatorname{ess\,sup}_{s \in [0,1]} f(s) - \frac{1}{\log n} \log \int_0^1 n^{f(s)} \, ds \leq \epsilon - \frac{\log \delta}{\log n}.$$

Of course the lower bound above is obvious. Now recall the modulus of continuity ω' from Lemma 5.3, and note that by its definition and that of Y_n we have $\omega'(Y_n, \delta) \leq \omega'(Z_n) + \delta$. Then for fixed $\epsilon > 0$ and $\delta \in (0, \epsilon/2)$ we have, by (5.9), that

$$\mathbb{P}\left(\operatorname{ess\,sup}_{s \in [0,1]} Y_n(s) - \frac{1}{\log n} \log \int_0^1 n^{Y_n(s)} \, ds > \epsilon - \frac{\log \delta}{\log n}\right) \leq \mathbb{P}\left(\sup_{|s-t|<\delta} |Y_n(s) - Y_n(t)| > \epsilon\right)$$

$$\leq \mathbb{P}(\omega'(Y_n, \delta) > \epsilon)$$

$$\leq \mathbb{P}(\omega'(Z_n, \delta) > \epsilon/2).$$

Therefore, for all n sufficiently large such that $-\log \delta / \log n < \epsilon$,

$$\frac{1}{K_n} \log \mathbb{P}\left(\operatorname*{ess\,sup}_{s \in [0,1]} Y_n(s) - \frac{1}{\log n} \log \int_0^1 n^{Y_n(s)} \, \mathrm{d}s > 2\epsilon\right) \leq \frac{1}{K_n} \log \mathbb{P}(\omega'(Z_n, \delta) > \epsilon/2).$$

Then take limsup of both sides as $n \to \infty$ and use that

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{K_n} \log \mathbb{P}(\omega'(Z_n, \delta) > \epsilon/2) = 0,$$

as proved in (5.5) of Lemma 5.3 to conclude. $\quad\square$

REFERENCES

[1] ARU, J., POWELL, E. and SEPÚLVEDA, A. (2019). Critical Liouville measure as a limit of subcritical measures. *Electron. Commun. Probab.* **24** Paper No. 18, 16. MR3933042 https://doi.org/10.1214/19-ECP209

[2] ASTALA, K., JONES, P., KUPIAINEN, A. and SAKSMAN, E. (2011). Random conformal weldings. *Acta Math.* **207** 203–254. MR2892610 https://doi.org/10.1007/s11511-012-0069-3

[3] BERESTYCKI, N. (2017). An elementary approach to Gaussian multiplicative chaos. *Electron. Commun. Probab.* **22** Paper No. 27, 12. MR3652040 https://doi.org/10.1214/17-ECP58

[4] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. Wiley, New York. MR1700749 https://doi.org/10.1002/9780470316962

[5] BOURGADE, P., NIKEGHBALI, A. and ROUAULT, A. (2009). Circular Jacobi ensembles and deformed Verblunsky coefficients. *Int. Math. Res. Not. IMRN* **23** 4357–4394. MR2558335 https://doi.org/10.1093/imrn/rnp092

[6] BREUER, J. (2010). Spectral and dynamical properties of certain random Jacobi matrices with growing parameters. *Trans. Amer. Math. Soc.* **362** 3161–3182. MR2592951 https://doi.org/10.1090/S0002-9947-10-04856-7

[7] BREUER, J., FORRESTER, P. J. and SMILANSKY, U. (2007). Random discrete Schrödinger operators from random matrix theory. *J. Phys. A*: *Math. Theor.* **40** F161–F168. https://doi.org/10.1007/978-3-662-06400-9

[8] BREUER, J., SIMON, B. and ZEITOUNI, O. (2018). Large deviations and sum rules for spectral theory: A pedagogical approach. *J. Spectr. Theory* **8** 1551–1581. MR3870076 https://doi.org/10.4171/JST/235

[9] BREUER, J., SIMON, B. and ZEITOUNI, O. (2018). Large deviations and the Lukic conjecture. *Duke Math. J.* **167** 2857–2902. MR3865654 https://doi.org/10.1215/00127094-2018-0027

[10] CANTERO, M. J., MORAL, L. and VELÁZQUEZ, L. (2003). Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl.* **362** 29–56. MR1955452 https://doi.org/10.1016/S0024-3795(02)00457-3

[11] CARDY, J. (2004). Calogero–Sutherland model and bulk-boundary correlations in conformal field theory. *Phys. Lett. B* **582** 121–126. https://doi.org/10.1016/j.physletb.2003.12.029

[12] CHHAIBI, R., MADAULE, T. and NAJNUDEL, J. (2018). On the maximum of the CβE field. *Duke Math. J.* **167** 2243–2345. MR3848391 https://doi.org/10.1215/00127094-2018-0016

[13] CHHAIBI, R. and NAJNUDEL, J. (2019). On the circle, $GMC^\gamma = \varprojlim C\beta E_n$ for $\gamma = \sqrt{\frac{2}{\beta}}$, ($\gamma \leq 1$). Available at arXiv:1904.00578.

[14] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. Springer, New York. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4

[15] DUPLANTIER, B. and SHEFFIELD, S. (2011). Liouville quantum gravity and KPZ. *Invent. Math.* **185** 333–393. MR2819163 https://doi.org/10.1007/s00222-010-0308-1

[16] DURRETT, R. (2010). *Probability*: *Theory and Examples*, 4th ed. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge Univ. Press, Cambridge. MR2722836

[17] DYSON, F. J. (1962). Statistical theory of the energy levels of complex systems. I. *J. Math. Phys.* **3** 140–156. MR0143556

[18] FENG, J. and KURTZ, T. G. (2006). *Large Deviations for Stochastic Processes*. *Mathematical Surveys and Monographs* **131**. Amer. Math. Soc., Providence, RI. MR2260560 https://doi.org/10.1090/surv/131

[19] FYODOROV, Y. V. and BOUCHAUD, J.-P. (2008). Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A*: *Math. Theor.* **41** 372001.

[20] GAMBOA, F., NAGEL, J. and ROUAULT, A. (2016). Sum rules via large deviations. *J. Funct. Anal.* **270** 509–559. MR3425894 https://doi.org/10.1016/j.jfa.2015.08.009

[21] GAMBOA, F., NAGEL, J. and ROUAULT, A. (2017). Sum rules and large deviations for spectral measures on the unit circle. *Random Matrices Theory Appl.* **6** 1750005, 49. MR3612269 https://doi.org/10.1142/S2010326317500058

[22] GAMBOA, F., NAGEL, J. and ROUAULT, A. (2019). Sum rules and large deviations for spectral matrix measures. *Bernoulli* **25** 712–741. MR3892334 https://doi.org/10.3150/17-bej1003

[23] GOOD, I. J. (1970). Short proof of a conjecture by Dyson. *J. Math. Phys.* **11** 1884. MR0258644

[24] HU, X., MILLER, J. and PERES, Y. (2010). Thick points of the Gaussian free field. *Ann. Probab.* **38** 896–926. MR2642894 https://doi.org/10.1214/09-AOP498

[25] JITOMIRSKAYA, S. and LAST, Y. (1999). Power-law subordinacy and singular spectra. I. Half-line operators. *Acta Math.* **183** 171–189. MR1738043 https://doi.org/10.1007/BF02392827

[26] KALLENBERG, O. (1986). *Random Measures*, 4th ed. Akademie-Verlag, Berlin; Academic Press, London. MR0854102

[27] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[28] KILLIP, R. and NENCIU, I. (2004). Matrix models for circular ensembles. *Int. Math. Res. Not.* **50** 2665–2701. MR2127367 https://doi.org/10.1155/S1073792804141597

[29] KILLIP, R. and NENCIU, I. (2007). CMV: The unitary analogue of Jacobi matrices. *Comm. Pure Appl. Math.* **60** 1148–1188. MR2330626 https://doi.org/10.1002/cpa.20160

[30] KILLIP, R. and RYCKMAN, E. (2010). Autocorrelations of the characteristic polynomial of a random matrix under microscopic scaling. Available online at arXiv:1004.1623.

[31] KILLIP, R. and STOICIU, M. (2009). Eigenvalue statistics for CMV matrices: From Poisson to clock via random matrix ensembles. *Duke Math. J.* **146** 361–399. MR2484278 https://doi.org/10.1215/00127094-2009-001

[32] LAST, Y. (1996). Quantum dynamics and decompositions of singular continuous spectra. *J. Funct. Anal.* **142** 406–445. MR1423040 https://doi.org/10.1006/jfan.1996.0155

[33] MILLER, J. and SHEFFIELD, S. (2016). Quantum Loewner evolution. *Duke Math. J.* **165** 3241–3378. MR3572845 https://doi.org/10.1215/00127094-3627096

[34] NIKULA, M., SAKSMAN, E. and WEBB, C. (2020). Multiplicative chaos and the characteristic polynomial of the CUE: The L^1-phase. *Trans. Amer. Math. Soc.* **373** 3905–3965. MR4105514 https://doi.org/10.1090/tran/8020

[35] REMY, G. (2020). The Fyodorov–Bouchaud formula and Liouville conformal field theory. *Duke Math. J.* **169** 177–211. MR4047550 https://doi.org/10.1215/00127094-2019-0045

[36] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1725357

[37] ROGERS, C. A. and TAYLOR, S. J. (1959). The analysis of additive set functions in Euclidean space. *Acta Math.* **101** 273–302. MR0107690

[38] ROGERS, C. A. and TAYLOR, S. J. (1963). Additive set functions in Euclidean space. II. *Acta Math.* **109** 207–240. MR0160860

[39] SIMON, B. (2005). *Orthogonal Polynomials on the Unit Circle*. *Part* 1: *Classical Theory*. American Mathematical Society Colloquium Publications **54**. Amer. Math. Soc., Providence, RI. MR2105088

[40] SIMON, B. (2005). *Orthogonal Polynomials on the Unit Circle*. *Part* 2: *Spectral Theory*. American Mathematical Society Colloquium Publications **54**. Amer. Math. Soc., Providence, RI. MR2105089

[41] VIRAG, B. (2018). Operator limits of random matrices. In *Proceedings of the International Congress of Mathematicians* **4**, Seoul.

[42] WEBB, C. (2015). The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos—the L^2-phase. *Electron. J. Probab.* **20** no. 104, 21. MR3407221 https://doi.org/10.1214/EJP.v20-4296

[43] WILLIAMS, D. (1991). *Probability with Martingales*. *Cambridge Mathematical Textbooks*. Cambridge Univ. Press, Cambridge. MR1155402 https://doi.org/10.1017/CBO9780511813658

[44] WILSON, K. G. (1962). Proof of a conjecture by Dyson. *J. Math. Phys.* **3** 1040–1043. MR0144627