

# REAL ROOTS OF RANDOM ORTHOGONAL POLYNOMIALS WITH EXPONENTIAL WEIGHTS

YEN DO, DORON LUBINSKY, HOI H. NGUYEN, OANH NGUYEN, AND IGOR PRITSKER

ABSTRACT. We consider random orthonormal polynomials

$$P_n(x) = \sum_{i=0}^n \xi_i p_i(x),$$

where  $\xi_0, \dots, \xi_n$  are independent random variables with zero mean, unit variance and uniformly bounded  $(2+\varepsilon_0)$ -moments, and  $\{p_n\}_{n=0}^\infty$  is the system of orthonormal polynomials with respect to a general exponential weight  $W$  on the real line. This class of orthogonal polynomials includes the popular Hermite and Freud polynomials. We establish universality for the leading asymptotics of the expected number of real roots of  $P_n$ , both globally and locally. In addition, we find an almost sure limit of the measures counting all roots of  $P_n$ . This is accomplished by introducing new ideas on applications of the inverse Littlewood-Offord theory in the context of the classical three term recurrence relation for orthogonal polynomials to establish anti-concentration properties, and by adapting the universality methods to the weighted random orthogonal polynomials of the form  $WP_n$ .

## 1. INTRODUCTION AND MAIN RESULTS

The study of roots of random polynomials has a long history which can be traced back to the fundamental papers by Bloch and Polya in 1932 [4], Littlewood-Offord [23, 24, 25], and Kac [19] in the 1940s, Erdős-Offord [11] in 1956, the list goes on. Given a basis of deterministic functions  $p_0, p_1, \dots$ , consider random polynomials of the form

$$P_n(x) = \xi_0 p_0(x) + \xi_1 p_1(x) + \dots + \xi_n p_n(x)$$

where  $\xi_i$  are independent random variables. Three popular classes of random polynomials include

- Random algebraic polynomials, which correspond to  $p_n = a_n x^n$ ,
- Random trigonometric polynomials, which correspond to  $p_n = a_n \sin(nx)$  and  $p_n = a_n \cos(nx)$ ,
- Random orthogonal polynomials, which correspond to an orthonormal basis of polynomials  $p_n$  associated with a given measure.

A very incomplete list of references for the first two classes include [1, 2, 3, 4, 8, 11, 13, 16, 17, 18, 19, 23, 24, 25, 31, 32, 36, 38]. In this paper, we focus on random orthogonal polynomials. Let  $\mu$  be a finite positive measure on the *real line* with finite moments of all orders, and with the associated orthonormal system of polynomials  $\{p_n\}_{n=0}^\infty$ . In other words,

each  $p_n$  has degree  $n$ , real coefficients, and a positive leading coefficient, and furthermore

$$\int_{\mathbb{R}} p_n(x)p_m(x)d\mu(x) = \begin{cases} 1, & m = n; \\ 0, & m \neq n; \end{cases}$$

see, e.g., Freud [14] for details. Assume furthermore that  $\mu$  is absolutely continuous with respect to the Lebesgue measure, that is  $d\mu(x) = w(x)dx$ , where  $w$  is referred to as the weight function. Some popular systems of orthogonal polynomials are

(E1) Legendre polynomials: The weight  $w(x)$  is the indicator function of  $[-1, 1]$ .

(E2) Chebyshev polynomials of type 1:  $w(x) = \frac{1}{\sqrt{1-x^2}}$  for  $x \in (-1, 1)$ , and zero elsewhere.

(E3) Chebyshev polynomials of type 2:  $w(x) = \sqrt{1-x^2}$  for  $x \in (-1, 1)$ , and zero elsewhere.

(E4) Jacobi polynomials: These are generalizations of Chebyshev and Legendre polynomials, with  $w(x) = (1-x)^\alpha(1+x)^\beta$  for  $x \in (-1, 1)$  and zero elsewhere, where  $\alpha, \beta > -1$ .

(E5) Hermite polynomials:  $w(x) = e^{-x^2}$  for  $x \in \mathbb{R}$ .

(E6) Freud polynomials: These are generalizations of Hermite polynomials, with  $w(x) = e^{-c|x|^\lambda}$  for  $x \in \mathbb{R}$ , where  $c > 0, \lambda > 0$ .

One of the most basic questions in the study of roots is the average number of real roots of  $P_n$ . In [19], Kac derived the celebrated Kac-Rice formula that allows us to reduce this task to evaluating an integral, which is often most manageable when the random variables  $\xi_i$  are standard Gaussian. Let  $N_n(S)$  be the number of roots of  $P_n$  in a set  $S$ . In this Gaussian setting, the average number of real roots of  $P_n$  has been known for the following models.

- Random Legendre polynomials: Das [5] found that  $\mathbb{E}[N_n(-1, 1)]$  is asymptotically equal to  $n/\sqrt{3}$ . Wilkins [39] improved the error term in this asymptotic relation by showing that  $\mathbb{E}[N_n(-1, 1)] = n/\sqrt{3} + o(n^\varepsilon)$  for any  $\varepsilon > 0$ ;
- Random Jacobi polynomials: Das and Bhatt [6] concluded that  $\mathbb{E}[N_n(-1, 1)]$  is also asymptotically equal to  $n/\sqrt{3}$ ;
- More generally, when  $\mu$  a finite Borel measure with *compact support* on the real line. Lubinsky–Pritsker–Xie [27] showed that under mild conditions on the weight  $d\mu/dx$ , the same asymptotic holds.
- Beyond compactly supported measures, Pritsker–Xie [33] extended this asymptotic to the random Freud polynomials. Later, Lubinsky–Pritsker–Xie [28] showed that the same asymptotic holds for general exponential weights (see Theorem 1.2 below).

Little is known about the average number of roots when the coefficients are not Gaussian. In [10], Do–Nguyen–Vu extended the asymptotics for the model in [27] to general random variables with mean 0, variance 1, and bounded  $(2 + \varepsilon_0)$ -moments for some  $\varepsilon_0 > 0$ . In other words, under some conditions, the asymptotic of  $\mathbb{E}N_n(\mathbb{R})$  is known when the measures  $\mu$  are compactly supported, such as the measures in Examples (E1)–(E4) above.

This paper is devoted to random orthogonal polynomials associated with exponential weights on the whole real line, such as random polynomials spanned by Hermite and Freud polynomials mentioned in (E5) and (E6). In particular, we show that the results of [28] hold for general random coefficients  $\xi_i$ . We shall assume throughout the paper that the measure  $\mu$  is absolutely continuous, with  $d\mu(x) = W^2(x)dx$ , where the weight  $W$  belongs to the class  $\mathcal{F}(C^2)$  of exponential weights considered in [22] and [28].

**Definition 1.1** (Weight class  $\mathcal{F}(C^2)$ ). *Let  $W = e^{-Q}$ , where  $Q : \mathbb{R} \rightarrow [0, \infty)$  satisfies the following conditions:*

- (a)  $Q'$  is continuous in  $\mathbb{R}$  and  $Q(0) = 0$ .
- (b)  $Q'$  is non-decreasing in  $\mathbb{R}$ , and  $Q''$  exists in  $\mathbb{R} \setminus \{0\}$ .
- (c)

$$\lim_{|t| \rightarrow \infty} Q(t) = \infty.$$

- (d) *The function*

$$T(t) = \frac{tQ'(t)}{Q(t)}, \quad t \neq 0,$$

*is quasi-increasing in  $(0, \infty)$ , in the sense that for some  $C_1 > 0$ ,*

$$0 < x < y \Rightarrow T(x) \leq C_1 T(y).$$

*We assume an analogous restriction for  $y < x < 0$ . In addition, we assume that for some  $\Lambda > 1$ ,*

$$T(t) \geq \Lambda \text{ in } \mathbb{R} \setminus \{0\}.$$

- (e) *There exists  $C_2 > 0$  such that*

$$\frac{Q''(x)}{|Q'(x)|} \leq C_2 \frac{|Q'(x)|}{Q(x)}, \quad x \in \mathbb{R} \setminus \{0\}.$$

*Then we write  $W \in \mathcal{F}(C^2)$ .*

It is straightforward to check that the Freud weights  $w(x) = e^{-c|x|^\lambda}$  of (E6) belong to  $\mathcal{F}(C^2)$  when  $\lambda > 1$ . From the definition,  $\mathcal{F}(C^2)$  consists of smooth functions  $e^{-Q}$ , where  $Q$  is convex, and has regular growth behavior at infinity. We note that  $\mathcal{F}(C^2)$  is one of the largest known classes of weights that allow for a very complete theory of the associated orthogonal polynomials, see [22] for a comprehensive exposition.

We recall the result of [28] on the asymptotic for  $\mathbb{E}N_n(\mathbb{R})$  where  $N_n(\mathbb{R})$  is the number of real roots of the random polynomial

$$P_n(x) = \sum_{i=0}^n \xi_i p_i(x).$$

**Theorem 1.2** (Gaussian setting, cf. Theorem 2.2 of [28]). *Let  $W = e^{-Q} \in \mathcal{F}(C^2)$ , where  $Q$  is even. If the function  $T$  in the definition of  $\mathcal{F}(C^2)$  satisfies*

$$(1.1) \quad \lim_{x \rightarrow \infty} T(x) = \alpha,$$

where  $\alpha \in (1, \infty]$ , then the expected number of real zeros of  $P_n$ , with  $\{\xi_i\}_{i=0}^\infty$  being i.i.d. standard Gaussian, satisfies

$$(1.2) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[N_n(\mathbb{R})] = \frac{1}{\sqrt{3}}.$$

To state the result on the local asymptotics for the expected number of real zeros, we first define the Ullman distribution that provides the limiting density for the number of real roots. Following [22] and [35], we define the Ullman distribution  $\mu_\alpha$ ,  $\alpha > 1$ , by

$$(1.3) \quad d\mu_\alpha(x) = \left( \frac{\alpha}{\pi} \int_{|x|}^1 \frac{t^{\alpha-1}}{\sqrt{t^2 - x^2}} dt \right) dx, \quad x \in [-1, 1],$$

and we consider the following contracted version of  $P_n$

$$(1.4) \quad P_n^*(s) := P_n(a_n s), \quad n \in \mathbb{N},$$

where  $a_n$  is the Mhaskar-Rakhmanov-Saff number associated with the weight  $W$ , see [22] and [35].

Roughly speaking, most of the real roots of  $P_n^*$  stay in  $[-1, 1]$ . For any set  $E \subset \mathbb{C}$ ,  $N_n^*(E)$  denotes the number of zeros of a random polynomial  $P_n^*$  located in  $E$ . We now state the local result on the asymptotic of  $\mathbb{E}[N_n^*([a, b])]$  for intervals  $[a, b] \subset (-1, 1)$ . It says that  $\mu_\alpha$  is the limiting measure, up to a constant.

**Theorem 1.3** (Gaussian setting, see Theorem 2.3 in [28]). *Assume that  $\{\xi_k\}_{k=0}^n$  are i.i.d. standard Gaussian. Let  $W = e^{-Q} \in \mathcal{F}(C^2)$ , where  $Q$  is even. Suppose that the function  $T$  in the definition of  $\mathcal{F}(C^2)$  satisfies (1.1). For any  $[a, b] \subset (-1, 1)$ , we have*

$$(1.5) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[N_n^*([a, b])] = \frac{1}{\sqrt{3}} \mu_\alpha([a, b]).$$

It is worth noting that condition (1.1) plays a crucial role in the existence of the limit in (1.5). In fact, if the weight function  $W$  does not behave at infinity in a regular way, by for example oscillating between the values of  $e^{-c|x|^{\lambda_1}}$  and  $e^{-c|x|^{\lambda_2}}$  for  $x$  near infinity, with  $\lambda_1 > \lambda_2 > 1$ , then the limit in (1.5) may fail to exist. This indicates that the local asymptotic behavior of  $\mathbb{E}[N_n^*([a, b])]$  may not abide a single limiting distribution, and may differ dramatically when the interval  $[a, b]$  is shifted.

We are now ready to state our main universality result confirming that Theorems 1.2 and 1.3 hold for general random coefficients. Throughout the paper, we assume that  $\xi_k$  are real-valued random variables.

**Theorem 1.4.** *Assume that the coefficients  $\{\xi_k\}_{k=0}^n$  are independent random variables such that  $\mathbb{E}[\xi_k] = 0$ ,  $\text{Var}[\xi_k] = 1$ , and  $\mathbb{E}[|\xi_k|^{2+\varepsilon_0}] < C$  for some constants  $C, \varepsilon_0 > 0$  and all  $k = 0, 1, \dots$ . Let  $W = e^{-Q} \in \mathcal{F}(C^2)$ , where  $Q$  is even. If the function  $T$  in the definition of  $\mathcal{F}(C^2)$  satisfies (1.1) with  $\alpha \in (1, \infty)$ , then we have*

- (a) *The expected number of real zeros of random orthogonal polynomial  $P_n$  satisfies (1.2).*
- (b) *For any  $[a, b] \subset (-1, 1)$ , the real zeros of polynomials  $P_n^*$  satisfy (1.5).*

In fact we will show that the  $k$ -correlations of the real roots of  $P_n^*$  are universal for any fixed  $k$ , see Theorem 2.2.

As a direct application, we obtain the statement for random Freud polynomials.

**Corollary 1.5.** *Assume that the coefficients  $\{\xi_k\}_{k=0}^n$  are independent random variables such that  $\mathbb{E}[\xi_k] = 0$ ,  $\text{Var}[\xi_k] = 1$ , and  $\mathbb{E}[|\xi_k|^{2+\varepsilon_0}] < C$  for some constants  $C, \varepsilon_0 > 0$  and all  $k = 0, 1, \dots$ . For any constants  $\lambda > 1$  and  $c > 0$ , consider the random orthogonal polynomial  $P_n$  associated with the Freud weight  $w(x) = e^{-c|x|^\lambda}$  for  $x \in \mathbb{R}$ . Then the expected number of real zeros of  $P_n$  satisfies (1.2). Moreover, if  $[a, b] \subset (-1, 1)$  then the real zeros of polynomials  $P_n^*$ , defined in (1.4), satisfy (1.5) with  $\alpha = \lambda$ .*

Next, define the normalized zero counting measure  $\tau_n = \frac{1}{n} \sum_{k=1}^n \delta_{z_{k,n}}$  for the scaled polynomial  $P_n^*$  (1.4), where  $\{z_{k,n}\}_{k=1}^n$  are its zeros in  $\mathbb{C}$ , counted with multiplicities, and  $\delta_z$  denotes the unit point mass at  $z$ . We find the weak\* limit of  $\tau_n$  in the case of random coefficients satisfying the assumptions of Theorem 1.4, and thus extend Theorem 2.4 of [28].

**Theorem 1.6.** *Assume that the coefficients  $\{\xi_k\}_{k=0}^n$  are independent random variables such that  $\mathbb{E}[\xi_k] = 0$ ,  $\text{Var}[\xi_k] = 1$ , and  $\mathbb{E}[|\xi_k|^{2+\varepsilon_0}] < C$  for some constants  $C, \varepsilon_0 > 0$  and all  $k = 0, 1, \dots$ . Let  $W = e^{-Q} \in \mathcal{F}(C^2)$ , where  $Q$  is even. If the function  $T$  in the definition of  $\mathcal{F}(C^2)$  satisfies (1.1) with  $\alpha \in (1, \infty)$ , then the normalized zero counting measures  $\tau_n$  for the scaled polynomials  $P_n^*(s)$  converge to  $\mu_\alpha$  of (1.3) in the weak\* topology with probability one.*

The following corollary is useful in the proof of Theorem 1.4.

**Corollary 1.7.** *Suppose that the assumptions of Theorem 1.6 hold. If  $E \subset \mathbb{C}$  is any compact set satisfying  $\mu_\alpha(\partial_{\mathbb{C}} E) = 0$ , then*

$$(1.6) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[N_n^*(E)] = \mu_\alpha(E).$$

**Remark 1.8.** In Theorem 1.6 and Corollary 1.7, the conditions on the random coefficients can be relaxed. In the proof for these results, we only need to assume that the random variables  $\{\xi_k\}_{k=0}^n$  are independent, moreover, there is a constant  $A > 0$  such that

$$\mathbb{E}[\log^+ |\xi_k|] < A, \quad k = 0, 1, \dots,$$

and there is a constant  $a \in (0, 1)$  such that

$$\mathbb{P}(|\xi_k - x| \leq a) \leq 1 - a, \quad k = 0, 1, \dots$$

## 2. THE METHOD OF PROOF AND UNIVERSALITY OF THE CORRELATIONS

To establish Theorem 1.4, we apply the universality framework of [31]. This approach was earlier developed in Tao–Vu [38], Do–Nguyen–Vu [9], and it now becomes one of the standard methods in the field. The global idea is to compare the distribution of roots of  $P_n^*$  with that of the Gaussian counterpart, namely

$$\tilde{P}_n^*(x) = \sum_{k=0}^n \tilde{\xi}_k p_k(a_n x)$$

where  $\tilde{\xi}_k$  are iid standard Gaussian. Generally speaking, if none of the  $p_k$  dominates the sum (see the Delocalization Condition C3 below), then the effect of changing the distribution of the coefficients is small and so the roots of  $P_n^*$  have roughly the same distribution as those of  $\tilde{P}_n^*$ . For technical reasons, and in particular to tame the growth of the orthogonal polynomials  $p_k$ , we shall work with the random functions

$$F_n(x) = P_n^*(x)W(a_n x) = \sum_{k=0}^n \xi_k p_k(a_n x)W(a_n x) = \sum_{k=0}^n \xi_k f_k(x),$$

where  $f_k(x) = p_k(a_n x)W(a_n x)$ ,  $k = 1, \dots, n$ . Since  $W(x) \neq 0$ ,  $x \in \mathbb{R}$ , the real zeros of  $F_n$  and  $P_n^*$  are clearly identical. Further, we need an extension of the weight  $W$  into the complex plane, which is convenient to obtain via the potential theoretic argument below. Following [22] and [35], we introduce the Mhaskar-Rakhmanov-Saff interval  $\Delta_n := [-a_n, a_n]$ . The equilibrium density (or the density of the weighted equilibrium measure) is defined as

$$\sigma_n(x) = \frac{\sqrt{(x - a_{-n})(a_n - x)}}{\pi^2} \int_{a_{-n}}^{a_n} \frac{Q'(s) - Q'(x)}{s - x} \frac{ds}{\sqrt{(s - a_{-n})(a_n - s)}}, \quad x \in \Delta_n.$$

It satisfies the equilibrium equation [22, p. 41]:

$$(2.1) \quad \int_{a_{-n}}^{a_n} \log \frac{1}{|x - t|} \sigma_n(t) dt + Q(x) = M, \quad x \in \Delta_n.$$

Note that the weighted equilibrium measure  $\sigma_n(x) dx$  is supported on  $\Delta_n$ , and has total mass  $n$ :

$$\int_{a_{-n}}^{a_n} \sigma_n(x) dx = n.$$

For details on  $\sigma_n$ , we refer to the book [22]. We now set

$$(2.2) \quad Q(z) := M - U^{\mu_n}(z) \text{ and } W(z) := \exp(-Q(z)), \quad z \in \mathbb{C} \setminus [-a_n, a_n],$$

where  $U^{\mu_n}(x)$  is the logarithmic potential of  $\mu_n$ :

$$U^{\mu_n}(x) := - \int_{a_{-n}}^{a_n} \log |x - t| \sigma_n(t) dt.$$

This gives the desired extension of  $W$  into  $\mathbb{C}$ , with  $Q$  being harmonic in  $\mathbb{C} \setminus [-a_n, a_n]$  and continuous in  $\mathbb{C}$ , cf. [22] and [35].

The exact conditions for the applicability of the universality method are stated below. For any constant  $\varepsilon > 0$ , we shall show that there exists a constant  $b > 0$  such that for any positive constants  $c_1, A$ , there exists  $C_1 > 0$  for which the following conditions hold for the interval  $D(\varepsilon) = [-1 + \varepsilon, 1 - \varepsilon]$ .

- (C1) *Anti-concentration*: For every  $z \in D(\varepsilon)$ , with probability at least  $1 - C_1 n^{-A}$ , there exists  $w \in B_{\mathbb{C}}(z, 1/n)$  for which  $|F_n(w)| \geq \exp(-n^{c_1})$ .
- (C2) *Boundedness*: For any  $z \in D(\varepsilon)$ , with probability at least  $1 - C_1 n^{-A}$ ,  $|F_n(w)| \leq \exp(n^{c_1})$  for all  $w \in B_{\mathbb{C}}(z, 1/n)$ .

(C3) *Delocalization*: For every  $z \in D(\varepsilon) + B_{\mathbb{C}}(0, 1/n)$ , for every  $k = 0, \dots, n$ ,

$$\frac{|p_k(a_n z)|}{\sqrt{\sum_{j=0}^n |p_j(a_n z)|^2}} \leq C_1 n^{-b}.$$

(C4) *Derivative growth*: For any  $s \in D(\varepsilon)$ ,

$$\sum_{j=0}^n |p'_j(a_n s)|^2 a_n^2 \leq C_1 n^{2+c_1} \sum_{j=0}^n |p_j(a_n s)|^2,$$

$$\sup_{z \in B_{\mathbb{C}}(s, 1/n)} \sum_{j=0}^n |p''_j(a_n z)|^2 a_n^4 \leq C_1 n^{4+c_1} \sum_{j=0}^n |p_j(a_n s)|^2.$$

**Lemma 2.1.** *Assume the hypothesis of Theorem 1.4. For any constant  $\varepsilon > 0$ , there exists a constant  $b > 0$  such that for any positive constants  $c_1, A$ , there exists  $C_1 > 0$  for which Conditions (C1)–(C4) hold for the interval  $D(\varepsilon) = [-1 + \varepsilon, 1 - \varepsilon]$ .*

Assuming Lemma 2.1, we have the following universal result of the correlations of real roots.

**Theorem 2.2.** *Assume the hypothesis of Theorem 1.4. Let  $k$  be any positive integer. Then there exist positive constants  $C$  and  $c$  independent of  $n$  such that the following holds. For any real numbers  $x_1, \dots, x_k$ , all of which are in  $D(\varepsilon)$ , and any function  $G : \mathbb{R}^k \rightarrow \mathbb{R}$  supported on  $\prod_{i=1}^k [x_i - c/n, x_i + c/n]$  with continuous derivatives up to order  $2k + 4$  and  $\|\nabla^m G\|_{\infty} \leq n^a$  for all  $0 \leq m \leq 2k + 4$ , we have*

$$\left| \mathbb{E} \left[ \sum G(\zeta_{i_1}, \dots, \zeta_{i_k}) \right] - \mathbb{E} \left[ \sum G(\tilde{\zeta}_{i_1}, \dots, \tilde{\zeta}_{i_k}) \right] \right| \leq C n^{-c},$$

where the first sum runs over all  $k$ -tuples  $(\zeta_{i_1}, \dots, \zeta_{i_k}) \in \mathbb{R}^k$  of the roots  $\zeta_1, \zeta_2, \dots$  of  $P_n^*$ , and the second sum runs over all  $k$ -tuples  $(\tilde{\zeta}_{i_1}, \dots, \tilde{\zeta}_{i_k}) \in \mathbb{R}^k$  of the roots  $\tilde{\zeta}_1, \tilde{\zeta}_2, \dots$  of  $\tilde{P}_n^*$ .

This theorem is a generalization of [31, Theorem 2.6]. We provide a proof in Section 9.

It is worth mentioning, by triangle inequality that Theorem 2.2 also holds true if  $\xi_k$  and  $\tilde{\xi}_k$  are any random variables with mean 0, variance 1, and uniformly bounded  $(2 + \varepsilon_0)$  moments. In other words, it is not necessary that  $\tilde{\xi}_k$  are standard Gaussian. **In Section 12 we will provide various formulas for the  $k$ -correlation function of the real roots of  $P_n^*$  (or of  $P_n$ ) under the assumption that the random coefficients  $\xi_i$  have smooth density functions.**

♣ **Hoi:** [To Oanh: I was thinking about a slightly more general version that  $\mathbb{E}\xi_i = \mathbb{E}\xi'_i$  and  $\mathbb{E}\xi_i^2 = \mathbb{E}(\xi'_i)^2$  (i.e. each  $\xi_i$  has matching moment up to order 2 with the corresponding  $\xi'_i$ , i.e. perturbed variant) and they have uniformly bounded  $(2 + \varepsilon_0)$  moments, would that work with minimal change (we might assume iid or just independent)? For instance we would like to apply to Exponential random variables, then they don't have mean zero. Also, I am assuming the  $\xi_i, \xi'_i$  are all real-valued here, right? I guess your comparison method works for complex as well, but if the coefficients are purely complex then it is unlikely that  $P_n(x)$  has real roots.]

Using Theorem 2.2, we derive an estimate for the number of real roots at a local scale, which immediately gives the second part of Theorem 1.4 by decomposing a big interval into

intervals of length  $O(1/n)$ . It is convenient here to use the notation  $N_{P_n^*}(E)$  and  $N_{\tilde{P}_n^*}(E)$  for the number of roots of  $P_n^*$  and  $\tilde{P}_n^*$  respectively in a set  $E$ .

**Theorem 2.3.** *Under the hypothesis of Theorem 1.4, for any  $\varepsilon > 0$ , there exists a positive constant  $c$  such that the following holds for any interval  $[a, b]$  of length  $O(1/n)$  inside  $D(\varepsilon)$ :*

$$\mathbb{E}[N_{P_n^*}([a, b])] - \mathbb{E}[N_{\tilde{P}_n^*}([a, b])] \ll n^{-c}.$$

Outside of the set  $D(\varepsilon)$ , we show that the number of roots is negligible.

**Theorem 2.4.** *Under the hypotheses of Theorem 1.4, we have*

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N_{P_n^*}(\mathbb{R} \setminus D(\varepsilon))]}{n} = \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N_{\tilde{P}_n^*}(\mathbb{R} \setminus D(\varepsilon))]}{n} = 0.$$

Combining the last two theorems, we obtain that

$$\lim_{n \rightarrow \infty} \left( \frac{\mathbb{E}[N_{P_n^*}(\mathbb{R})]}{n} - \frac{\mathbb{E}[N_{\tilde{P}_n^*}(\mathbb{R})]}{n} \right) = 0.$$

This proves the first part of Theorem 1.4, in view of Theorem 1.2.

### 3. PROOF IDEAS AND OUR CONTRIBUTIONS

As mentioned in the previous section, the global structure of our paper is to follow the universality framework that has been developed in [31, 38], and more recently in [10] for random orthogonal polynomials with compactly supported measures with a “nice” density function. However, handling exponential weights is a much more complicated task that requires new tools and ideas.

Firstly, the function  $F_n$  whose universality property is of interest were assumed to be analytic in previous papers, which was important for several of their arguments such as the one using Halász’s inequality. Unfortunately, our function  $F_n(x) = P_n(x)W(x)$  is not analytic. We propose a novel way to circumvent this condition in Section 9.

Secondly, the framework of universality requires to investigate the behavior of  $F_n$  in an open neighborhood of the real line in the complex plane. That means we need to extend and study the behavior of the weight function  $W$  on  $\mathbb{C}$ . This brings technical challenges that one did not have to deal with in the previous papers such as [28].

Last but not least, one of the most challenging obstacles that we have to face is to show that the function  $F_n(x)$  cannot be too small on any interval  $I$  of length  $1/n$  inside the core interval  $(-a_n, a_n)$ ; this property is known as the anti-concentration property. If  $F_n(x)$  were too small, the function would become sensitive to noise, and one would not expect universality. So, it is crucial to establish the anti-concentration. In particular, we want to show that with high probability, there exists  $x \in I$  such that

$$|F_n(x)| \geq \exp(-n^\varepsilon)$$

for some small  $\varepsilon > 0$ .

When  $I$  is very near the edge of  $(-a_n, a_n)$  (that is  $x = \Theta(a_n)$ ),  $W(x)$  can be as small as  $\exp(-\Theta(n))$  and hence the above inequality becomes

$$(3.1) \quad |P_n(x)| \geq \exp(\Theta(n^\varepsilon)).$$

This is in sharp contrast to the case when the measure  $\mu$  is compact in which one only needs to show

$$|P_n(x)| \geq \exp(-n^\varepsilon).$$

The difficulty here is that with the power  $\varepsilon$  for  $n$ , one can restrict to lower degree polynomials  $p_k$  for  $k \ll n^{2\varepsilon}$  and use classical tools like Remez inequality. This was one of the key ingredients in both [31] and [10]. Now, having the power 1 in (3.1), one can no longer disregard the high degree polynomials and use Remez inequality as before. It boils down to controlling  $p_k(x)$  for  $k \approx n$  and  $|x| \approx a_n$ . This is known to have erratic behavior and it would be reasonable to expect to find  $I$  near  $a_n$  on which  $p_k$  can be very tiny.

In this paper, we come up with a completely new argument that incorporates the inverse Littlewood-Offord theory, the polynomial recurrence structure of  $p_k$ , and Remez inequality. We note that recurrence was also used in [38, Section 13] but our implementation here is completely orthogonal. The general idea is that if  $P_n$  has a small anti-concentration probability, then  $p_k$  must exhibit an algebraic structure. We then show that it is impossible for  $p_k$  to have both algebraic structure and polynomial recurrence structure. We refer the details to Section 7, where we also include a sketch of proof.

Finally, we obtain the  $k$ -correlation function formulas (in Section 12) by adapting the method of [15] with some minimal modifications toward orthogonal polynomials.

**Organization.** To prove Lemma 2.1, in Sections 4–7, we verify Conditions (C1)–(C4). In Section 8, we prove Theorem 2.3. Section 10 contains a proof of Theorem 1.6. Theorem 2.4 is proven in Section 11.

**Notations.** We use standard asymptotic notations under the assumption that  $n$  tends to infinity. For two positive sequences  $(a_n)$  and  $(b_n)$ , we say that  $a_n \gg b_n$  or  $b_n \ll a_n$  if there exists a constant  $C$  such that  $b_n \leq Ca_n$ . If  $|c_n| \ll a_n$  for some sequence  $(c_n)$ , we also write  $c_n \ll a_n$ .

If  $a_n \ll b_n \ll a_n$ , we say that  $b_n = \Theta(a_n)$ . If  $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 0$ , we say that  $a_n = o(b_n)$ . If  $b_n \ll a_n$ , we sometimes employ the notations  $b_n = O(a_n)$  and  $a_n = \Omega(b_n)$  to make the idea intuitively clearer or the writing less cumbersome; for example, if  $A$  is the quantity of interest, we may write  $A = A' + O(B)$  instead of  $A - A' \ll B$ , and  $A = e^{O(B)}$  instead of  $\log A \ll B$ . We also write that  $a_n = O_C(b_n)$  if the implied constant depends on a given parameter  $C$ .

For the orthonormal polynomials  $\{p_j(x)\}_{j=0}^\infty$ , we define the **reproducing kernel** by

$$K_n(x, y) = \sum_{j=0}^n p_j(x)p_j(y),$$

and the differentiated kernels by

$$K_n^{(k,l)}(x, y) = \sum_{j=0}^n p_j^{(k)}(x)p_j^{(l)}(y), \quad k, l \in \mathbb{N} \cup \{0\}.$$

It is apparent that the above Conditions (C1)–(C4) can be rephrased in term of reproducing kernels.

#### 4. BOUNDEDNESS

In this section, we prove Condition (C2) as the second part of Lemma 2.1. Let

$$J_n(\varepsilon) = a_n D(\varepsilon) = [-(1 - \varepsilon)a_n, (1 - \varepsilon)a_n].$$

Since  $P_n$  is spanned by the orthonormal basis  $\{p_k\}_{k=0}^n$ , we have that

$$\|WP_n\|_{L^2(\mathbb{R})} = \left( \int_{-\infty}^{\infty} |P_n(x)|^2 W^2(x) dx \right)^{1/2} = \left( \sum_{k=0}^n |\xi_k|^2 \right)^{1/2}.$$

Applying the Nikolskii-type inequality (cf. [28, p. 60]), valid for any polynomial  $Q_n$  with complex coefficients of degree at most  $n$ ,

$$\|WQ_n\|_{L^\infty(\mathbb{R})} \leq Cn \|WQ_n\|_{L^2(\mathbb{R})}, \quad n \in \mathbb{N},$$

we obtain that

$$(4.1) \quad \|WP_n\|_{L^\infty(\mathbb{R})} \leq Cn \left( \sum_{k=0}^n |\xi_k|^2 \right)^{1/2}, \quad n \in \mathbb{N}.$$

It follows from Lemma 4.1 below (with  $m = 0$ ) that there is a constant  $B > 0$  such that for all  $n \in \mathbb{N}$ ,  $s \in D(\varepsilon) = [-1 + \varepsilon, 1 - \varepsilon]$ , we have

$$(4.2) \quad W(a_n s) \sup_{z \in B_{\mathbb{C}}(s, 1/n)} |P_n(a_n z)| \leq B \|WP_n\|_{L^\infty(\mathbb{R})} \leq Cn \left( \sum_{k=0}^n |\xi_k|^2 \right)^{1/2}.$$

From Chebyshev's inequality for random coefficients with finite  $(2 + \varepsilon_0)$ -moments, we obtain that  $\mathbb{P}(|\xi_k| > n^{(A+1)/2}) = O(n^{-(A+1)})$  holds for any  $A > 0$ ,  $k = 0, 1, \dots, n$ . Hence

$$\left( \sum_{k=0}^n |\xi_k|^2 \right)^{1/2} \leq n^{(A+1)/2} \sqrt{n+1}$$

and

$$W(a_n s) \sup_{z \in B_{\mathbb{C}}(s, 1/n)} |P_n(a_n z)| \leq Cn^{2+A/2}, \quad n \in \mathbb{N}, \quad s \in D(\varepsilon),$$

with probability at least  $(1 - cn^{-(A+1)})^{n+1} \geq 1 - d/n^A$ ,  $d > 0$ . Since

$$\sup_{z \in B_{\mathbb{C}}(s, 1/n)} W(a_n z) \leq CW(a_n s)$$

by our choice of the extension for  $W$  to the complex plane, we have that

$$\sup_{z \in B_{\mathbb{C}}(s, 1/n)} |P_n(a_n z)| W(a_n z) \leq \sup_{z \in B_{\mathbb{C}}(s, 1/n)} |P_n(a_n z)| \sup_{z \in B_{\mathbb{C}}(s, 1/n)} W(a_n z) \leq Cn^{2+A/2}$$

for all  $n \in \mathbb{N}$ ,  $s \in D(\varepsilon)$ , with probability at least  $1 - d/n^A$ ,  $d > 0$ . This verifies the boundedness condition (2) because  $W(a_n z)P_n(a_n z) = F_n(z)$  and  $n^{2+A/2} = o(\exp(c_1 n))$  for all  $c_1 > 0$ .

We now state and prove the following lemma about growth of weighted polynomials and their derivatives in the complex plane. It was already used in the above proof, namely in (4.2), and will also be very useful in the other sections.

**Lemma 4.1.** *For all  $n \in \mathbb{N}$ ,  $s \in D_n(\varepsilon)$ , and polynomials  $R_n$ ,  $\deg(R_n) \leq n$ , there is a constant  $B > 0$  such that*

$$W(a_n s) \sup_{z \in B_{\mathbb{C}}(s, 1/n)} |R_n^{(m)}(a_n z)| \leq B \left( \frac{n}{a_n} \right)^m \sup_{t \in [-1, 1]} |W(a_n t) R_n(a_n t)|, \quad m = 0, 1, 2, \dots$$

*Proof of Lemma 4.1.* We begin with a general estimate for the growth of an arbitrary polynomial  $R_n$ ,  $\deg(R_n) \leq n$ , in the complex plane in terms of its weighted supremum norm on  $\Delta_n = [-a_n, a_n]$ . Recall that the potential  $U^{\mu_n}(x) := -\int_{-a_n}^{a_n} \log|x-t| \sigma_n(t) dt$  of the weighted equilibrium measure  $\mu_n(x) := \sigma_n(x) dx$  satisfies the equilibrium equation from (2.1)

$$U^{\mu_n}(x) + Q(x) = M, \quad x \in \Delta_n.$$

It is known that  $U^{\mu_n}(x)$  is superharmonic and continuous in  $\mathbb{C}$ , harmonic in  $\mathbb{C} \setminus \Delta_n$ , and behaves like  $-n \log|x|$  near infinity. Therefore the function

$$h(z) := \log |R_n(z)| - \log \|R_n W\|_{L^\infty(\Delta_n)} + U^{\mu_n}(z) - M$$

is subharmonic in the domain  $\overline{\mathbb{C}} \setminus \Delta_n$ , and has the following boundary values

$$\begin{aligned} h(x) &= \log |R_n(x)| - \log \|R_n W\|_{L^\infty(\Delta_n)} - Q(x) \\ &= \log |R_n(x) W(x)| - \log \|R_n W\|_{L^\infty(\Delta_n)} \leq 0, \quad x \in \Delta_n. \end{aligned}$$

Applying the Maximum Principle for subharmonic functions, we conclude that  $h(z) \leq 0$  for all  $z \in \overline{\mathbb{C}} \setminus \Delta_n$ , which is equivalent to

$$(4.3) \quad |R_n(z)| \leq \|R_n W\|_{L^\infty(\Delta_n)} \exp(M - U^{\mu_n}(z)), \quad z \in \mathbb{C}.$$

This upper bound allows us to estimate the growth of  $R_n$ , provided we have appropriate estimates on the potential  $U^{\mu_n}(z)$ . Required estimates follow from Lemma 5.10(a) of [22, p. 130]: There is a constant  $C > 0$  such that

$$(4.4) \quad U^{\mu_n}(x + iy) - U^{\mu_n}(x) = O(1)$$

holds uniformly for  $n \geq n_0$ ,  $x \in J_n(\varepsilon)$ , and  $|y| \leq Ca_n/n$ . Furthermore, for any pair  $x_1, x_2 \in \Delta_n$  such that  $|x_1 - x_2| \leq ca_n/n$ , we obtain that

$$(4.5) \quad |U^{\mu_n}(x_1) - U^{\mu_n}(x_2)| = |Q(x_1) - Q(x_2)| \leq |x_1 - x_2| \max_{x \in \Delta_n} |Q'(x)| \leq c \frac{a_n Q'(a_n)}{n} = O(1),$$

where we used that  $a_n Q'(a_n) = O(n)$  by (3.7) and (3.11) of [28]. The latter estimate together with (4.4) implies that

$$U^{\mu_n}(z) = U^{\mu_n}(x) + O(1) = -Q(x) + O(1), \quad z \in B_{\mathbb{C}}(x, ca_n/n), \quad x \in J_n(\varepsilon),$$

where  $c > 0$  is independent of  $n, x, z$ . Hence (4.3) now gives

$$(4.6) \quad \begin{aligned} |R_n(z)| &\leq C \|R_n W\|_{L^\infty(\Delta_n)} \exp(M - U^{\mu_n}(x)) \\ &= C \frac{\|R_n W\|_{L^\infty(\Delta_n)}}{W(x)}, \quad z \in B_{\mathbb{C}}(x, ca_n/n), \quad x \in J_n(\varepsilon). \end{aligned}$$

Differentiating Cauchy's integral formula, we deduce from (4.6) that

$$(4.7) \quad \begin{aligned} |R_n^{(m)}(z)| &= \left| \frac{m!}{2\pi i} \int_{|w-x|=2a_n/n} \frac{R_n(w) dw}{(w-z)^{m+1}} \right| \\ &\leq C \left( \frac{n}{a_n} \right)^m \frac{\|R_n W\|_{L^\infty(\Delta_n)}}{W(x)}, \quad z \in B_{\mathbb{C}}(x, a_n/n), \quad x \in J_n(\varepsilon). \end{aligned}$$

The statement of Lemma 4.1 now follows by changing variable  $x = a_n s$ , while passing from  $\Delta_n$  to  $[-1, 1]$  and from  $J_n(\varepsilon)$  to  $D_n(\varepsilon)$ .  $\square$

## 5. DELOCALIZATION

In this section, we prove Condition (C3) as the third part of Lemma 2.1.

Recall that  $K_n(x, x) = \sum_{j=0}^n p_j^2(x)$ , so that  $|K_n(x, x)| = K_n(x, x)$ ,  $x \in \mathbb{R}$ , and  $|K_n(w, w)| \leq \sum_{j=0}^n |p_j(w)|^2$ ,  $w \in \mathbb{C}$ . We also use the notation  $J_n(\varepsilon) = a_n D(\varepsilon) = [-(1-\varepsilon)a_n, (1-\varepsilon)a_n]$ . In order to prove the delocalization condition, we write

$$(5.1) \quad \frac{|p_k(w)|}{\sqrt{\sum_{j=0}^n |p_j(w)|^2}} \leq \frac{|p_k(w)|}{\sqrt{|K_n(w, w)|}} = \frac{|p_k(w)W(x)|}{\sqrt{K_n(x, x)W^2(x)}} \left( \frac{K_n(x, x)}{|K_n(w, w)|} \right)^{1/2},$$

where we assume that  $x \in J_n(\varepsilon)$  and  $z \in \mathbb{C}$ . Our first goal is to find upper bounds for the orthonormal polynomials  $p_k$ . Applying Theorem 1.18 of [22, p. 22], we obtain the estimate

$$\sup_{t \in [-1, 1]} |W(a_n t) p_k(a_n t)| \leq \sup_{x \in \mathbb{R}} |W(x) p_k(x)| \leq C k^{1/6} a_k^{-1/3} (T(a_k)/a_k)^{1/6},$$

that holds uniformly for all  $k \geq 1$ . Since  $T$  is bounded as a continuous function on  $\mathbb{R}$  with finite limit at  $\infty$ , cf. (1.1), we have that

$$(5.2) \quad \sup_{t \in [-1, 1]} |W(a_n t) p_k(a_n t)| \leq \sup_{x \in \mathbb{R}} |W(x) p_k(x)| = O\left(k^{1/6} a_k^{-1/2}\right), \quad k \in \mathbb{N}.$$

Combining the above estimate with Lemma 4.1 (for  $m = 0$ ), we arrive at

$$(5.3) \quad W(a_n s) \sup_{z \in B_{\mathbb{C}}(s, 1/n)} |p_k(a_n z)| = O\left(k^{1/6} a_k^{-1/2}\right), \quad k \in \mathbb{N},$$

which holds uniformly for  $s \in D(\varepsilon)$ .

We now recall some facts about the reproducing kernel  $K_n$ . Theorem 1.25 of [22, p. 26] states that

$$\lim_{n \rightarrow \infty} K_n(x, x)W^2(x) = \sigma_n(x), \quad x \in J_n(\varepsilon),$$

where convergence is uniform in  $J_n(\varepsilon)$ . Further, Lemma 5.1(a) of [20, p. 87] gives that uniformly for  $x \in J_n(\varepsilon)$ ,

$$C_1 \frac{n}{a_n} \leq \sigma_n(x) \leq C_2 \frac{n}{a_n}.$$

Hence we have that

$$(5.4) \quad C_1 \frac{n}{a_n} \leq K_n(x, x)W^2(x) \leq C_2 \frac{n}{a_n}, \quad x \in J_n(\varepsilon),$$

which implies (after changing variable  $x = a_n s$ ) by (5.3) that

$$(5.5) \quad \frac{|p_k(a_n z)|W(a_n s)}{\sqrt{K_n(a_n s, a_n s)W^2(a_n s)}} \leq C \frac{k^{1/6}}{n^{1/2}} \left( \frac{a_n}{a_k} \right)^{1/2}, \quad s \in D(\varepsilon), z \in B_{\mathbb{C}}(s, 1/n),$$

where the estimate holds uniformly in  $z, s$  and  $k \leq n \in \mathbb{N}$ . Suppose first that  $k \leq n^\tau$ ,  $\tau \in (0, 1)$ . Recall that  $a_n$  is an increasing sequence of positive numbers, and that Lemma 3.5(c) of [22, p. 72] states for a constant  $C > 0$ :

$$(5.6) \quad 1 \leq \frac{a_n}{a_1} \leq Cn^{1/\Lambda} \text{ for all } n \in \mathbb{N},$$

where  $\Lambda > 1$  is from the Definition 1.1 of  $W$ . Hence (5.5) gives that

$$(5.7) \quad \frac{|p_k(a_n z)|W(a_n s)}{\sqrt{K_n(a_n s, a_n s)W^2(a_n s)}} \leq Cn^{(1/\Lambda-1)/2+\tau/6}, \quad s \in D(\varepsilon), z \in B_{\mathbb{C}}(s, 1/n),$$

uniformly for  $k \leq n^\tau$ . If  $n^\tau \leq k \leq n$  then we use that  $K_n(x, x) \geq K_k(x, x)$ , together with (5.3) and (5.4), and write instead of (5.5) that

$$(5.8) \quad \frac{|p_k(a_n z)|W(a_n s)}{\sqrt{K_n(a_n s, a_n s)W^2(a_n s)}} \leq \frac{|p_k(a_n z)|W(a_n s)}{\sqrt{K_k(a_n s, a_n s)W^2(a_n s)}} \\ \leq C \frac{k^{1/6}}{k^{1/2}} = Ck^{-1/3} \leq Cn^{-\tau/3}, \quad s \in D(\varepsilon), z \in B_{\mathbb{C}}(s, 1/n).$$

Selecting the value of  $\tau > 0$  sufficiently small, so that  $b := \min(\tau/3, (1 - 1/\Lambda)/2 - \tau/6) > 0$ , we obtain from (5.7)-(5.8) that

$$(5.9) \quad \frac{|p_k(a_n z)|W(a_n s)}{\sqrt{K_n(a_n s, a_n s)W^2(a_n s)}} \leq Cn^{-b}, \quad s \in D(\varepsilon), z \in B_{\mathbb{C}}(s, 1/n),$$

holds uniformly for all  $k \leq n$ .

Lemma 8.3(a) of [26] states that uniformly for  $u, v$  in compact subsets of the complex plane, and  $x \in J_n(\varepsilon)$ , we have as  $n \rightarrow \infty$ ,

$$(5.10) \quad \lim_{n \rightarrow \infty} \frac{K_n \left( x + \frac{u}{\bar{K}_n(x, x)}, x + \frac{v}{\bar{K}_n(x, x)} \right)}{K_n(x, x)} e^{-\frac{Q'(x)}{\bar{K}_n(x, x)}(u+v)} = \frac{\sin \pi(v - u)}{\pi(v - u)},$$

where  $\tilde{K}_n(x, x) := K_n(x, x)W^2(x)$ . The original result (5.10) for the real parameters  $u, v$  was established in Theorem 1.2 of [21]. Recall that

$$|Q'(x)| \leq C \frac{n}{a_n}$$

uniformly in  $x \in J_n(\varepsilon)$ , by Lemma 3.8(a) of [22, p. 77]. Combining the latter fact with (5.4), we obtain that

$$(5.11) \quad 0 < C_1 \leq \frac{Q'(x)}{\tilde{K}_n(x, x)} \leq C_2$$

holds uniformly for  $x \in J_n(\varepsilon)$ . Thus (5.10) now gives that

$$\frac{K_n(x, x)}{|K_n(w, w)|} \leq C$$

uniformly for  $x \in J_n(\varepsilon)$  and  $w \in B_{\mathbb{C}}(x, a_n/n)$ . If we change variables  $x = a_n s$  and  $w = a_n z$  in the above estimate, and use the result together with (5.9) in (5.1), then we arrive at the estimate

$$(5.12) \quad \frac{|p_k(a_n z)|}{\sqrt{\sum_{j=0}^n |p_j(a_n z)|^2}} \leq C n^{-b}, \quad k = 0, 1, \dots, n,$$

uniformly for  $s \in D(\varepsilon)$  and  $z \in B_{\mathbb{C}}(s, 1/n)$ .

## 6. DERIVATIVE GROWTH

In this section, we prove Condition (C4) as the last part of Lemma 2.1. Both inequalities of (C4) are obtained from the universality limit for the reproducing kernels stated in (5.10). Note that the functions under the limit on the right of (5.10) are entire functions in each complex variable  $u$  and  $v$  that converge uniformly in  $u$  and  $v$  on compact subsets of  $\mathbb{C}$ , and also uniformly for  $x \in J_n(\varepsilon) = [-(1-\varepsilon)a_n, (1-\varepsilon)a_n]$ . Hence the same uniform convergence on compact subsets of  $\mathbb{C}$  is preserved for all derivatives of these functions with respect to  $u$  and  $v$ . Differentiating (5.10) and applying induction, we obtain that

$$(6.1) \quad \lim_{n \rightarrow \infty} \frac{K_n^{(l,m)} \left( x + \frac{u}{\tilde{K}_n(x,x)}, x + \frac{v}{\tilde{K}_n(x,x)} \right)}{(\tilde{K}_n(x,x))^{l+m} K_n(x,x)} e^{-\frac{Q'(x)}{\tilde{K}_n(x,x)}(u+v)} = F_{l,m}(u, v, x),$$

where  $F_{l,m}(u, v, x)$  are entire in  $u$  and  $v$ , and the convergence is uniform for  $u$  and  $v$  in compact subsets of  $\mathbb{C}$ , and for  $x \in J_n(\varepsilon)$ . It follows that each function  $F_{l,m}(u, v, x)$  is uniformly bounded for  $u$  and  $v$  in compact subsets of  $\mathbb{C}$ , and for  $x \in J_n(\varepsilon)$ . Taking into account (5.11), we obtain that

$$\frac{K_n^{(l,m)} \left( x + \frac{u}{\tilde{K}_n(x,x)}, x + \frac{v}{\tilde{K}_n(x,x)} \right)}{(\tilde{K}_n(x,x))^{l+m} K_n(x,x)}$$

are bounded uniformly for all  $n \in \mathbb{N}$ ,  $u$  and  $v$  in compact subsets of  $\mathbb{C}$ , and  $x \in J_n(\varepsilon)$ . Recall that  $\tilde{K}_n(x, x) = K_n(x, x)W^2(x)$  satisfies (5.4), which gives that

$$(6.2) \quad \frac{K_n^{(l,m)} \left( x + \frac{ua_n}{n}, x + \frac{va_n}{n} \right)}{K_n(x, x)} \leq C_{l,m} \left( \frac{n}{a_n} \right)^{l+m},$$

uniformly for all  $n \in \mathbb{N}$ ,  $u$  and  $v$  in compact subsets of  $\mathbb{C}$ , and  $x \in J_n(\varepsilon)$ .

To check the first inequality in condition (C4), we use (6.2) with  $l = m = 1$ ,  $u = v = 0$ , and  $x = a_n s$ , where  $s \in D(\varepsilon)$ , and estimate

$$\frac{\sum_{j=0}^n |p'_j(a_n s)|^2 a_n^2}{\sum_{j=0}^n |p_j(a_n s)|^2} = a_n^2 \frac{K_n^{(1,1)}(a_n s, a_n s)}{K_n(a_n s, a_n s)} = O(n^2).$$

Thus the first inequality in condition (C4) is true with  $c_1 = 0$ .

We now turn to the proof of the second derivative growth condition, where we apply (6.2) with  $l = m = 2$ ,  $|u| < 1$  and  $|v| < 1$ , and  $x = a_n s$ , with  $s \in D(\varepsilon)$ . It follows that

$$\frac{\sup_{z \in B_{\mathbb{C}}(s, 1/n)} \sum_{j=0}^n |p''_j(a_n z)|^2 a_n^4}{\sum_{j=0}^n |p_j(a_n s)|^2} = a_n^4 \frac{\sup_{z \in B_{\mathbb{C}}(s, 1/n)} K_n^{(2,2)}(a_n z, a_n \bar{z})}{K_n(a_n s, a_n s)} = O(n^4).$$

We also used that the coefficients of the orthogonal polynomials  $p_j$  are real, to express the numerator in the above equation via the kernel  $K_n^{(2,2)}$ . This completes verification of the derivative growth conditions with  $c_1 = 0$ .

## 7. ANTI-CONCENTRATION

In this section, we prove Condition (C1) as a part of Lemma 2.1. In particular, we show that the minimum of  $P_n^*$  in a local ball is bounded away from 0 with high probability. The key result of this section is

**Lemma 7.1** (Anti-concentration). *For any positive constants  $A, c$  and  $c_1$ , the following holds: for every  $x_0 \in [-1 + \varepsilon, 1 - \varepsilon]$ , with probability at least  $1 - O(n^{-A})$ , there exists  $x_1 \in [x_0 - c/n, x_0 + c/n]$  for which  $|W(a_n x_1) P_n^*(x_1)| \geq \exp(-n^{c_1})$ .*

We will prove an actually stronger result below

**Lemma 7.2.** *For any positive constants  $A, c$  and  $c_1$ , the following holds: for every  $x_0 \in [-1 + \varepsilon, 1 - \varepsilon]$ , there exists  $x_1 \in [x_0 - c/n, x_0 + c/n]$  such that*

$$\mathbb{P}(|W(a_n x_1) P_n^*(x_1)| \leq \exp(-n^{c_1})) < n^{-A}.$$

The rest of this section is devoted to the proof of this lemma. Before going into the detail, let us first provide a road map.

- (1) Step 0: Assume by contradiction that Lemma 7.2 fails.
- (2) Step 1: We recall the classical three-term recurrence for three consecutive  $p_m$ . From that, we can write  $p_{m+k}$  in terms of  $p_m$  and  $p_{m-1}$  which we call a jumped three-term recurrence (see Lemma 7.4).
- (3) Step 2: From the inverse Littlewood-Offord theory, we know that for a given  $x$ , the function  $W(a_n x) P_n^*(x) = \sum_{i=0}^n \xi_i W(a_n x) p_i(a_n x)$  to satisfy

$$\mathbb{P}(|W(a_n x) P_n^*(x)| \leq \exp(-n^{c_1})) \geq n^{-A},$$

it is necessary that the components  $W(a_n x)p_i(a_n x)$  are close to a general arithmetic progression. Roughly speaking, this means that they have a low linear-dimension. See Lemma 7.5.

- (4) Step 3: From Step 2, we know that there are many non-trivial linear combinations of the  $(W(a_n x)p_i(a_n x))$  that are very small

$$\sum_{j=1}^{O(1)} W(a_n x)p_{i_j}(a_n x) \approx 0 \text{ (which precisely means } \exp(-n^c)\text{-close to 0).}$$

Using the three-term recurrence in Step 1 to write all of these  $p_{i_j}$  as combination of  $p_m$  and  $p_{m-1}$ , we obtain a so-called two-term degeneracy

$$W(a_n x)p_m(a_n x) \approx \Gamma(x)W(a_n x)p_{m-1}(a_n x)$$

for some low-degree fractional polynomial  $\Gamma$ . See Lemma 7.7.

- (5) Step 4: Adapting the same strategy from Step 3, we can reduce two-term degeneracy to (one-term) degeneracy. See Lemma 7.9.
- (6) Step 5: Show that degeneracy is impossible as it reduces to a polynomial of low degree and large leading term taking small value on a significant interval.

We shall now go into the details.

**7.1. Step 1: Jumped three-term recurrence.** The following lemma covers a classic fact concerning the three-term recurrence for  $(p_m)$ .

**Lemma 7.3.** *The polynomials  $p_m$  satisfy a three-term recurrence relation*

$$(7.1) \quad xp_m = A_m p_{m+1} + B_m p_m + A_{m-1} p_{m-1}$$

Moreover, there exists  $r \in (0, 1)$  and  $C_1, C_2 > 0$  such that for  $m \geq 1$ ,

$$\begin{aligned} C_1 &\leq A_m \leq C_2 m^r; \\ |B_m| &\leq C_2 m^r. \end{aligned}$$

For our purpose, we only need a weaker statement  $m^{-C} \ll A_m \ll m^C$  and  $B_m \ll m^C$ .

*Proof of Lemma 7.3.* We let  $a_{\pm m}$  denote the Mhaskar-Rakhmanov-Saff numbers, so that

$$\begin{aligned} m &= \frac{1}{\pi} \int_{a_{-m}}^{a_m} \frac{xQ'(x)}{\sqrt{(x-a_{-m})(a_m-x)}} dx; \\ 0 &= \frac{1}{\pi} \int_{a_{-m}}^{a_m} \frac{Q'(x)}{\sqrt{(x-a_{-m})(a_m-x)}} dx. \end{aligned}$$

We let

$$\begin{aligned} \delta_m &= \frac{1}{2} (a_m + |a_{-m}|); \\ \beta_m &= \frac{1}{2} (a_m + a_{-m}), \end{aligned}$$

so that

$$[a_{-m}, a_m] = [\beta_m - \delta_m, \beta_m + \delta_m].$$

is the  $m$ th Mhaskar-Rakhmanov-Saff interval.

From [22, Theorem 15.2, p. 402], which holds for a larger class of weights,

$$\lim_{m \rightarrow \infty} \frac{A_m}{\delta_m} = \frac{1}{2} \text{ and } \lim_{m \rightarrow \infty} \frac{B_m - \beta_m}{\delta_m} = 0.$$

Note that by definition,  $|\beta_m| \leq \delta_m$ . So all we need to prove is that there exists  $r < 1$  such that both

$$\delta_m = O(m^r).$$

This follows from [22, Lemma 3.4(c), p. 72], where it is proved that there exists  $r_0$  such that for  $t \geq r \geq r_0$ ,

$$1 \leq \frac{\delta_t}{\delta_r} \leq C \left( \frac{t}{r} \right)^{1/\Lambda}.$$

Since  $\Lambda > 1$ , we can take  $r = 1/\Lambda$ . □

From the recurrence, we obtain that

$$\begin{aligned} p_{m+1} &= \left( \frac{1}{A_m} x + \dots \right) p_m + \frac{A_{m-1}}{A_m} p_{m-1}, \\ p_{m+2} &= \left( \frac{1}{A_{m+1}} x + \dots \right) p_{m+1} + \frac{A_{m-1}}{A_m} p_m \\ &= \left( \frac{1}{A_{m+1} A_m} x^2 + \dots \right) p_m + \left( \frac{A_{m-1}}{A_{m+1} A_m} x + \dots \right) p_{m-1}, \end{aligned}$$

and so, by the bounds from Lemma 7.3, we obtain the following ‘‘jump’’ recurrence.

**Lemma 7.4.** *For all  $k, m \geq 1$ ,*

$$(7.2) \quad p_{m+k} = \left( \frac{1}{\prod_{i=m}^{m+k-1} A_i} x^k + \dots \right) p_m + \left( \frac{A_{m-1}}{\prod_{i=m}^{m+k-1} A_i} x^{k-1} + \dots \right) p_{m-1},$$

where in the dots are the lower order terms whose coefficients are  $O(n^{O_k(1)})$ .

**7.2. Large anti-concentration probability implies arithmetic structure.** Let

$$\beta = \exp(-n^{c_1}).$$

Let  $x \in [x_0 - c/n, x_0 + c/n]$  and assume otherwise that  $\mathbb{P}(|W(a_n x) P_n^*(x)| \geq \beta) > n^{-A}$ . Then by [29, Theorem 2.9] (see also [30, 37]), there exists an exceptional set  $B_x \subset [n]$  of size at most  $n^{\delta/2}$  so that the set  $\{|W(a_n x) p_i(a_n x), i \notin B_x\}$  are  $\beta$ -closed to a GAP of size  $n^{O_A(1)}$  and rank at most  $r' = O_A(1)$ . In other words, there exist a set of  $r'$  generators  $g_1, \dots, g_{r'}$  (all depending on  $x$ ) such that for all  $i \notin B_x$ , there exists integers  $k_{i1}, \dots, k_{i r'}$ , all of order  $n^{O_A(1)}$  satisfying

$$(7.3) \quad |W(a_n x) p_i(a_n x) - \sum_{j=1}^{r'} k_{ij} g_j| = O(\beta).$$

In the following lemma, by passing to a subset, we can reduce to the case that the rank  $r'$  and the coefficients  $k_{ij}$  are the same. By the delocalization bound in Section 5, we can also reduce to a subset of induces  $n'$  at which  $|W(a_n x)p_{n'}(a_n x)|$  is not too small.

**Lemma 7.5.** *There exist constants  $B, \delta > 0$  such that for any constant  $K$ , there exist a measurable set  $\mathcal{E} \subset [x_0 - c/n, x_0 + c/n]$  and an index  $n^\delta < n' < n - n^\delta$  such that*

- (1)  $\lambda_{Leb}(\mathcal{E}) > 2c/n^2$ ;
- (2)  $|W(a_n x)p_{n'}(a_n x)| \geq n^{-B}, x \in \mathcal{E}$ ;
- (3)  $[n', n' + K] \cap B_x = \emptyset, x \in \mathcal{E}$ .
- (4) *The GAP ranks  $r'$  are identical for all  $x \in \mathcal{E}$ . Moreover, the coefficients  $k_{ij}(x) = k_{ij}(y)$  for all  $n' \leq i \leq n' + K, 1 \leq j \leq r', x, y \in \mathcal{E}$ .*

Here,  $\lambda_{Leb}$  is the Lebesgue measure on  $\mathbb{R}$ .

In this section we will choose  $K$  sufficiently large compared to  $r$  and  $A$ , but still  $K = O(1)$ .

*Proof.* We start with the second item. For  $x \in [x_0 - c/n, x_0 + c/n]$ , by (5.12), for some constant  $b > 0$  and for all  $0 \leq i \leq n$ ,

$$\frac{W(a_n x)|p_i(a_n x)|}{W(a_n x)\sqrt{K_n(x, x)}} = \frac{W(a_n x)|p_i(a_n x)|}{W(a_n x)\sqrt{\sum_{j=0}^n p_j^2(a_n x)}} \leq Cn^{-b}.$$

Thus, for  $\delta = b/2$ , the contribution of  $i \in [n^\delta, n - n^\delta]$  is dominant. By (5.4) and (5.6), for some constant  $\Lambda > 1$ ,

$$K_n(x, x)W^2(a_n x) \gg \frac{n}{a_n} \gg \frac{n}{n^{1/\Lambda}} = n^{1-1/\Lambda}.$$

Thus,

$$\sum_{i=n^\delta}^{n-n^\delta} W(a_n x)|p_i(a_n x)| \geq \frac{1}{2}.$$

So, there exists  $B > 0$  sufficiently large so that for each  $x$ , the set  $N_x$  of indices  $i$  where  $n^\delta < i < n - n^\delta$  and  $|W(a_n x)p_i(a_n x)| > n^{-B}$  satisfies

$$|N_x| \geq n^\delta.$$

For each  $x \in [x_0 - c/n, x_0 + c/n]$ , because the set of  $i$  for which  $\{i, \dots, i+K\} \cap B_x \neq \emptyset$  has size at most  $2Kn^{\delta/2}$ , and because  $|N_x| \geq n^\delta$ , there exists  $i \in N_x$  for which  $\{i, \dots, i+K\} \cap B_x = \emptyset$ . For each such  $x$ , we fix such an index  $i = i(x)$  (note that by continuity, for  $x$  very close to  $x'$  we can choose  $i(x) = i(x')$ ). When  $x$  varies over  $[x_0 - c/n, x_0 + c/n]$ , by pigeonhole principle, there exists a measurable set  $\mathcal{E}'$  of measure  $\lambda_{Leb}(\mathcal{E}') \geq (2c/n)/n$  where  $i(x)$  takes the same value for all  $x \in \mathcal{E}'$ . Set  $n'$  to be this common value, we complete the proof.

Recall that by the application of [29, Theorem 2.9] above, we can identify  $W(a_n x)p_i(a_n x), i \in \{n', n' + 1, \dots, n' + K\}$  with a vector  $\mathbf{k}_i = \mathbf{k}_i(x) = (k_{i1}, \dots, k_{ir'}) \in \mathbb{Z}^{r'}$  with  $|k_{ir'}| \leq n^{O_A(1)}$

and  $r' = O(1)$ . More importantly, as there are only  $n^{O_A(1)}$  such choices of vectors for  $n' \leq i \leq n' + K$ , by passing to a subset  $\mathcal{E} \subset \mathcal{E}'$ , still of measure  $n^{-\Omega_{K,A}(1)}$ , we can assume that these vectors  $\mathbf{k}_i(x)$  are all the same for all  $x \in \mathcal{E}$ .  $\square$

In the next step, we show that there are many  $n''$  in the range  $n' \leq n'' \leq n' + K$  such that  $|W(a_n x)p_{n''}(a_n x)| \geq n^{-B}$ ,  $x \in \mathcal{E}$ .

**Claim 7.6.** Let  $B'$  be sufficiently large depending on  $B$  and  $K$ . Then for all  $0 \leq i \leq K$ , we must have either  $|W(a_n x)p_{n'+i}(a_n x)| \geq n^{-B'}$  or  $|W(a_n x)p_{n'+i+1}(a_n x)| \geq n^{-B'}$  for each  $x \in \mathcal{E}$ .

*Proof.* Assume otherwise, we can use the recurrence (7.1) to go backward and use the lower bound for  $A_m$  from Lemma 7.3 to arrive at  $|W(a_n x)p_{n'}(a_n x)| \leq n^{-B}$ , a contradiction.  $\square$

**7.3. Arithmetic structure + jumped three-term recurrence imply 2-term degeneracy.** The following lemma shows the so-called 2-term degeneracy of the form

$$W(a_n x)p_{n_0+1}(a_n x) \approx W(a_n x)p_{n_0}(a_n x)$$

for a large number of choices for  $n_0$ .

**Lemma 7.7.** Let  $K$  be an arbitrarily large constant and  $\varepsilon_1$  be a sufficiently small constant. Let  $n'$  be as in Lemma 7.5. There exists a sub-interval  $[n_1, n_2] \subset [n', n' + K]$  of length  $n_2 - n_1 \gg \log K$  such that for all  $n_0 \in [n_1, n_2]$ , we have

$$W(a_n x)p_{n_0+1}(a_n x) = -\frac{\frac{A_{n_0}}{\prod_{i=n_0+1}^{n_0+r-1} A_i} x^{r-2} + \dots}{\frac{1}{\prod_{i=n_0+1}^{n_0+r-1} A_i} x^{r-1} + \dots} W(a_n x)p_{n_0}(a_n x) + O(n^{O(1)} \beta^{1-\varepsilon_1}), \quad x \in \mathcal{E}.$$

Here, we note that the set  $\mathcal{E}$  may be a subset of the  $\mathcal{E}$  in Lemma 7.5 but it always has a Lebesgue measure of  $\Omega(n^{-2})$ .

*Proof of Lemma 7.7.* We consider the vectors  $\mathbf{k}_{n'}, \dots, \mathbf{k}_{n'+K}$ . For each  $n' \leq i \leq j \leq n' + K$ , we let

$$G_{[i,j]} = \{\mathbf{k}_i, \dots, \mathbf{k}_j\}$$

and  $H_{[i,j]}$  be the subspace (over  $\mathbb{R}$ ) spanned by the vectors in  $G_{[i,j]}$ . Set

$$s_0 = \lfloor (\log_2(K/2))/2 \rfloor.$$

For  $0 \leq s \leq s_0$ , consider the nested subsets  $G_{[n'+K/2-2^{s_0+s}, n'+K/2+2^{s_0+s}]}$ . As the dimensions  $\dim(H_{[n'+K/2-2^{s_0+s}, n'+K/2+2^{s_0+s}]})$  are non-decreasing in  $s$ , and all bounded by  $r'$ , by pigeonhole principle, provided that  $K$  is sufficiently large given  $r'$ , there exists  $s_1 \leq s_0 - 1$  such that

$$(7.4) \quad \dim(H_{[n'+K/2-2^{s_0+s_1}, n'+K/2+2^{s_0+s_1}]}) = \dim(H_{[n'+K/2-2^{s_0+s_1+1}, n'+K/2+2^{s_0+s_1+1}]})$$

Let  $n_1 = n' + K/2 - 2^{s_0+s_1+1}$ ,  $n'_1 = n' + K/2 + 2^{s_0+s_1+1}$ ,  $n_2 = n' + K/2 - 2^{s_0+s_1}$ , and  $n'_2 = n' + K/2 + 2^{s_0+s_1}$  be the endpoints.



We can take any small constant  $\varepsilon_1 \leq \frac{1}{2}$  that satisfies this lemma.

Taking union over all  $n^{O(1)}$  choices of  $Q_{n_0}$ , the set of points  $x$  at which  $|Q_{n_0}(x)| \leq \beta^{\varepsilon_1}$  has measure at most  $\beta^{2\varepsilon_1/r}$ . So, by removing these  $x$  from  $\mathcal{E}$ , the new set which we again denote by  $\mathcal{E}$  still has size  $n^{-\Omega(1)}$ . On this set, on (7.6), by dividing by  $Q_{n_0}(x)$ , we get

$$W(a_n x) p_{n_0+1}(a_n x) = -\frac{\frac{A_{n_0}}{\prod_{i=n_0+1}^{n_0+r-1} A_i} x^{r-2} + \dots}{\frac{1}{\prod_{i=n_0+1}^{n_0+r-1} A_i} x^{r-1} + \dots} W(a_n x) p_{n_0}(a_n x) + O(n^{O(1)} \beta^{1-\varepsilon_1}), \quad x \in \mathcal{E}.$$

This finishes the proof of Lemma 7.7.  $\square$

*Proof of Lemma 7.8.* For the first part, we have

$$|S(x)| \leq \sum_{i=0}^d |a_i| x^i \leq (d+1) n^B n^{3Bd}.$$

For the second part, let

$$E = \{x \in I_B : |S(x)| \leq \beta^{\varepsilon_1}\}.$$

To bound the size of  $E$ , we shall use the classical Remez inequality: for every measurable subset  $E \subset J$ , we have

$$\max_{I_B} |S| \leq \left( \frac{4|I_B|}{|E|} \right)^d \sup_E |S| \leq \left( \frac{4|I_B|}{|E|} \right)^d \beta^{\varepsilon_1}.$$

Since

$$\max_{I_B} |S| \geq |S(n^{3B})| \geq |a_d| n^{3Bd} - \sum_{i=0}^{d-1} |a_i| n^{3Bi} \geq n^{3Bd-B} - dn^{3B(d-1)+B} \geq 1$$

for sufficiently large  $n$ . Thus,

$$|E| \leq 8n^{3B} \beta^{\varepsilon_1/d}.$$

This completes the proof of Lemma 7.8.  $\square$

**7.4. Arithmetic structure + two-term degeneracy imply degeneracy.** We shall now derive a (one-term) degeneracy of the form

$$W(a_n x) p_{n'_1}(a_n x) \approx 0.$$

**Lemma 7.9.** *Let  $[n_1, n_2]$  be as in Lemma 7.7. There exists  $n_3 \in [n_1, n_2]$  such that*

$$\frac{R_1(x)}{R_2(x)} W(a_n x) p_{n_3}(a_n x) = O(\beta^{1-\varepsilon_1}), \quad x \in \mathcal{E}$$

where  $R_1$  and  $R_2$  are polynomials satisfying the hypothesis of Lemma 7.8 and  $\deg(R_1) = \deg(R_2) - 1$ .

*Proof of Lemma 7.9.* We continue using the notations in the proof of Lemma 7.7. Let  $r_0 = n_2 - n_1 - 1$ . Consider the vectors  $\mathbf{k}_{n_1+i}$ ,  $0 \leq i \leq r_0$ . Because the vectors are in  $\mathbb{R}^{r'}$  and  $r_0$  is much large than  $r'$ , there must exist  $k$  such that

$$\mathbf{k}_{n_1+2k}, \mathbf{k}_{n_1+2k+1} \in \langle \mathbf{k}_{n_1+2k+1+i}, 1 \leq i \leq r_0 - 2k - 1 \rangle.$$

On the other hand, we have learned from Claim 7.6 that either  $|W(a_n x)p_{n_1+2k}(a_n x)| \geq n^{-B'}$  or  $|W(a_n x)p_{n_1+2k_1}(a_n x)| \geq n^{-B'}$ . Assume without loss of generality that  $|W(a_n x)p_{n_1+2k}(a_n x)| \geq n^{-B'}$ . Set

$$n_3 = n_1 + 2k.$$

As  $\mathbf{k}_{n_3} \in \langle \mathbf{k}_{n_3+i}, 0 \leq i \leq r_0 - 2k \rangle$ , from (7.3), we can find rational coefficients  $d_\ell$  of height at most  $n^{O(1)}$  such that

$$(7.7) \quad W(a_n x)p_{n_3+1}(a_n x) - \sum_{\ell=2}^{r_0-2k} d_\ell W(a_n x)p_{n_3+\ell}(a_n x) = O(n^{O(1)}\beta), \quad x \in \mathcal{E}.$$

Recall from Lemma 7.7 that for all  $j = 1, \dots, r_0 - 1$ ,

$$(7.8) \quad W(a_n x)p_{n_1+j+1}(a_n x) = -\frac{\frac{A_{n_1+j}}{\prod_{i=n_1+j+1}^{n_1+j+r-1} A_i} x^{r-2} + \dots}{\frac{1}{\prod_{i=n_1+j+1}^{n_1+j+r-1} A_i} x^{r-1} + \dots} W(a_n x)p_{n_1+j}(a_n x) + O(n^{O(1)}\beta^{1-\varepsilon_1}).$$

Applying (7.8) for  $j = n_3 - n_1$ , and then recursively for  $j = n_3 - n_1 + 1, \dots, n_3 - n_1 + r_0 - 2k$  and all  $x \in \mathcal{E}$ ,

$$\begin{aligned} W(a_n x)p_{n_3+s+1}(a_n x) &= -\frac{\frac{\prod_{\ell=0}^s A_{n_3+\ell}}{\prod_{\ell=0}^s \prod_{i=n_3+\ell+1}^{n_3+\ell+r-1} A_i} x^{(s+1)(r-2)} + \dots}{\frac{1}{\prod_{\ell=0}^s \prod_{i=n_3+\ell+1}^{n_3+\ell+r-1} A_i} x^{(s+1)(r-1)} + \dots} W(a_n x)p_{n_3}(a_n x) + O(n^{O(1)}\beta^{1-\varepsilon_1}) \\ &=: -\frac{\bar{A}_{s+1} x^{(s+1)(r-2)} + \dots}{\hat{A}_{s+1} x^{(s+1)(r-1)} + \dots} W(a_n x)p_{n_3}(a_n x) + O(n^{O(1)}\beta^{1-\varepsilon_1}) \\ &=: \Gamma_{s+1}(x) W(a_n x)p_{n_3}(a_n x) + O(n^{O(1)}\beta^{1-\varepsilon_1}). \end{aligned}$$

Substitute into (7.7) we obtain

$$R(x)W(a_n x)p_{n_3}(a_n x) = O(n^{O(1)}\beta^{1-\varepsilon_1}), \quad x \in \mathcal{E}$$

where

$$R(x) = \Gamma_1(x) - \sum_{\ell=2}^{r_0-2k} d_\ell \Gamma_\ell(x).$$

Since  $\Gamma_\ell(x)$  is a fraction of two polynomials with the numerator's degree is  $\ell$  less than the denominator's degree,  $R(x)$  can be written as  $R(x) = \frac{R_1(x)}{R_2(x)}$  where  $R_2(x)$  is the product of all denominators of  $\Gamma_\ell(x)$  and  $\deg(R_1) = \deg(R_2) - 1$ . The leading coefficient of  $R_1$  is

$$\bar{A}_1 \hat{A}_2 \dots \hat{A}_{r_0-2k} \geq n^{-\Omega(1)}.$$

Moreover, all coefficients of  $R_1$  are  $O(n^{O(1)})$  by Lemma 7.3. So,  $R_1$  satisfies the hypothesis of Lemma 7.8. The same reasoning applies for  $R_2$ . This completes the proof of Lemma 7.9.  $\square$

7.5. **Degeneracy is impossible.** We are now ready to derive a contradiction which completes the proof of Lemma 7.2. Since  $|W(a_n x)p_{n_3}(a_n x)| \geq n^{-B}$  by the definition of  $\mathcal{E}'$ , we get

$$R(x) = O(n^{O(1)}\beta^{1-\varepsilon_1}), \quad x \in \mathcal{E}.$$

Since  $R_2$  satisfies the hypothesis of Lemma 7.8, we get that for all  $x \in \mathcal{E}$ ,

$$|R_2(x)| \geq n^{O(1)}\beta^{1-\varepsilon_1}.$$

So,

$$|R_1(x)| = O(n^{O(1)}\beta^{1-\varepsilon_1}), \quad x \in \mathcal{E}.$$

Since  $R_1$  also satisfies the hypothesis of Lemma 7.8 and since  $|\mathcal{E}| \geq n^{-O(1)}$  much larger than  $n^{O(1)}\beta^{1-\varepsilon_1}$ , we obtain a contradiction to Lemma 7.8.

## 8. PROOF OF THEOREM 2.3

Our idea is to apply Theorem 2.2 to the indicator function  $\mathbf{1}_{[a,b]}$ . However, since the indicator function is not smooth, we will approximate it below and above by smooth functions, say  $F_1$  and  $F_2$ . To this end, let  $\alpha = n^{-1-c/12}$  where  $c$  is the constant in Theorem 2.2 and let  $F_1, F_2 : \mathbb{R} \rightarrow \mathbb{R}$  be any continuous functions that are continuously differentiable up to degree 6 and satisfy the following

- $\mathbf{1}_{[a-\alpha, b+\alpha]} \geq F_2 \geq \mathbf{1}_{[a,b]} \geq F_1 \geq \mathbf{1}_{[a+\alpha, b-\alpha]}$  pointwise,
- $\|F_i^a\|_\infty \ll \alpha^{-a} \ll n^{a+c/2}$  for all  $i = 1, 2$  and  $1 \leq a \leq 6$ .

Applying Theorem 2.2 to  $G_i = n^{-c/2}F_i$  (where the factor  $n^{-c/2}$  is added so that  $\|G_i^a\|_\infty \ll n^{-a}$  as in the hypothesis), we obtain

$$\mathbb{E} \sum G_1(\zeta_i) - \mathbb{E} \sum G_1(\tilde{\zeta}_i) \ll n^{-c}, \quad \mathbb{E} \sum G_2(\zeta_i) - \mathbb{E} \sum G_2(\tilde{\zeta}_i) \ll n^{-c}$$

which gives

$$(8.1) \quad \mathbb{E} \sum F_1(\zeta_i) - \mathbb{E} \sum F_1(\tilde{\zeta}_i) \ll n^{-c/2}, \quad \mathbb{E} \sum F_2(\zeta_i) - \mathbb{E} \sum F_2(\tilde{\zeta}_i) \ll n^{-c/2}.$$

We shall prove later that

$$(8.2) \quad \mathbb{E} \sum \tilde{F}_1(\zeta_i) - \mathbb{E} \sum \tilde{F}_2(\zeta_i) \ll n^{-c/24}.$$

Since  $\sum F_1(\zeta_i) \leq N[a, b] \leq \sum F_2(\zeta_i)$ , we conclude from (8.1) and (8.2) that

$$\mathbb{E}N_{P_n^*}[a, b] - \mathbb{E}N_{\tilde{P}_n^*}[a, b] \ll n^{-c/24}$$

as claimed.

To prove (8.2), we note that  $0 \leq F_2 - F_1 \leq \mathbf{1}_{[a-\alpha, a+\alpha]} + \mathbf{1}_{[b-\alpha, b+\alpha]}$ , thus it suffices to show the following

**Lemma 8.1.** *Let  $I \subset D(\varepsilon)$  be an interval of length  $n^{-1-c/12}$ . Then*

$$\mathbb{E}N_{\tilde{P}_n^*}(I) \ll n^{-c/24}.$$

*Proof.* By the Kac-Rice formula [19], we have

$$\begin{aligned} \frac{1}{n} \mathbb{E} N_{\tilde{P}_n^*}(I) &= \frac{1}{n} \mathbb{E} N_{\tilde{P}_n}(a_n I) = \int_{a_n I} \frac{1}{\pi} \frac{K_n(x, x)}{n} \sqrt{\frac{K_n^{(1,1)}(x, x)}{K_n(x, x)^3} - \left( \frac{K_n^{(0,1)}(x, x)}{K_n(x, x)^2} \right)^2} dx \\ &\leq \int_{a_n I} \frac{1}{n\pi} \sqrt{\frac{K_n^{(1,1)}(x, x)}{K_n(x, x)}} dx = \int_{a_n I} \frac{1}{n\pi} \sqrt{\frac{K_n^{(1,1)}(x, x) W^2(x)}{K_n(x, x) W^2(x)}} dx. \end{aligned}$$

By (5.4), we have

$$\inf_{x \in [-1+\varepsilon, 1-\varepsilon]} K_n(x, x) W^2(x) \gg \frac{n}{a_n}.$$

By (5.4) and (6.2) (with  $u = v = 0$ ), we have

$$\sup_{x \in [-1+\varepsilon, 1-\varepsilon]} K_n^{(1,1)}(x, x) W^2(x) \ll \frac{n^3}{a_n^3}.$$

Therefore,

$$\frac{1}{n} \mathbb{E} N_{\tilde{P}_n^*}(I) \ll \int_{a_n I} \frac{1}{a_n} dx \ll |I| = n^{-1-c/12}$$

as desired.  $\square$

## 9. PROOF OF THEOREM 2.2

The proof roughly follows [31, Theorem 2.6] with our  $\mathbf{F}_n$  plays the role of the  $F_n$  in [31].

The only issue with our  $F_n$  is that it is not analytic. And there are two places in [31] that use analyticity. We shall explain how we walk around these two places.

In the first place, is to derive an alternative version of [31, Equation 11] where the use of the Green's formula is no longer applicable to our  $F_n$  because it is not necessarily analytic. Firstly, we observe that the roots of  $P_n^*$  and  $F_n$  are the same. Thus, for any smooth function  $G_j$  on  $\mathbb{C}$ , we get the following equations which can be used as a replacement of [31, Equation 11]

$$\begin{aligned} \sum_i G_j(\zeta_i) &= \int_{\mathbb{C}} \log |P_n^*(z)| H_j(z) dz \\ &= \int_{B(z_j, 1/10)} \log |F_n(u_j)| H_j(u_j) du_j - c(G), \end{aligned}$$

where  $H_j(z) = \frac{1}{2\pi} \Delta G_j(z)$  and

$$c(G) = \int_{B(z_j, 1/10)} \log |W(a_n u_j)| H_j(u_j) du_j$$

is a deterministic number.

The second place, also more crucial, is the treatment of [31, Lemma 8.2] which originally requires that  $F_n$  is analytic. Here, we restate the statement for our  $F_n$  which is not necessarily analytic.

**Lemma 9.1.** *Let  $0 < c_2 < 1$  and let  $F_n$  be a function of the form  $F_n(x) = P_n^*(x)W(a_n x)$  where  $P_n^*$  is an entire function and  $-\log W = Q$  satisfies (4.5). Assume that  $|F_n(w)| \geq \exp(-n^{c_2})$  for some  $w \in [-1 + \varepsilon, 1 - \varepsilon]$  and  $|F_n(z)| \leq \exp(n^{c_2})$  for all  $z \in B(w, 3/(2n))$ . Then*

$$\int_{B(w, 1/(2n))} |\log |F_n(z)||^2 dz \leq Cn^{-2+6c_2}.$$

*Proof of Lemma 9.1.* We follow ideas from [9]; the constant 6 in the conclusion is adhoc but we make no attempt to optimize it.

From Jensen's inequality for the number of roots of  $P_n^*$ , we have

$$N_{F_n}(B(w, 1/n)) = N_{P_n^*}(B(w, 1/n)) \leq \log \frac{5}{2} (\log M - \log |P_n^*(w)|) < 2(\log M - \log |P_n^*(w)|)$$

where  $N_{F_n}(B(w, 1/n))$  is the number of zeros of  $F_n$  in  $B(w, 1/n)$  and  $M = \max_{|w-z|=2/n} |P_n^*(z)|$ .

We have

$$\frac{\max_{|w-z|=2/n} |P_n^*(z)|}{|P_n^*(w)|} \leq \frac{\max_{|w-z|=2/n} |F_n(z)|}{|F_n(w)|} \sup_{|w-z|=2/n} \frac{W(a_n w)}{W(a_n z)} \ll \frac{\max_{|w-z|=2/n} |F_n(z)|}{|F_n(w)|}$$

where in the last inequality, we used (4.5).

From this and the assumption of the lemma, we conclude that

$$(9.1) \quad N_{F_n}(B(w, 1/n)) \leq 2n^{c_2}.$$

By the pigeonhole principle, there exists a radius  $1/n \geq r \geq 1/(2n)$  for which  $F_n$  has no zeros in the annulus  $B(w, r + \eta) \setminus B(w, r - \eta)$  where  $\eta = .1n^{-1-c_2}$ . We can also assume, without loss of generality, that there is no root on the boundary of each disk.

Let  $\zeta_1, \dots, \zeta_m$  be the zeros of  $F_n$  in the disk  $B(w, r - \eta)$ . By (9.1),  $m \leq 2n^{c_2}$ . Define

$$p(z) := \frac{P_n^*(z)}{(nz - n\zeta_1) \dots (nz - n\zeta_m)} \quad \text{and} \quad f(z) := p(z)W(a_n z).$$

Since  $p$  is an entire function which does not have zeros in the (closed) disk  $B(w, r + \eta)$ ,  $\log |p|$  is harmonic on this disk. For every  $z$  with  $|z - w| = r + \eta$ , the distance from  $z$  to any  $\zeta_i$  is at least  $\eta$ , so

$$|p(z)| \leq |P_n^*(z)| n^{-m} \eta^{-m} \ll \exp(n^{c_2}) (n\eta)^{-m} W^{-1}(a_n w) =: \exp(n^{c_2}) (n\eta)^{-m} \alpha^{-1}$$

where  $\alpha = W(a_n w)$  and we used  $\inf_{|z-w|=r+\eta} W(a_n z) = \Theta(W(a_n w))$  by (4.5).

It follows that for any  $z$  where  $|z - w| = r + \eta$

$$(9.2) \quad \log |p(z)| \leq n^{-c_2} + m \log(n\eta)^{-1} + \log \alpha + O(1) \leq 21n^{2c_2} - \log \alpha,$$

since

$$n^{c_2} \leq n^{2c_2}, m \leq 2n^{c_2}, (n\eta)^{-1} = 10n^{c_2} \leq e^{10n^{c_2}}.$$

Because of the harmonicity of  $\log |p|$ , its maximum is achieved on the boundary, and so the same bound holds for all  $z \in B(w, r + \eta)$ . On the other hand, from the lower bound on  $|F_n(w)|$  in the lemma and the fact that  $|\zeta_i - w| \leq 1/n$ ,

$$(9.3) \quad \log |p(w)| \geq \log |P_n^*(w)| \geq -n^{c_2} - \log \alpha.$$

Now, we make a critical use of Harnack's inequality [34, Chapter 11], which asserts that if a function  $G$  is harmonic on the open disk  $B(w, R)$  and is nonnegative continuous on its closure, for some  $w \in \mathbb{C}$  and  $R > 0$ , then for every  $z \in B(w, r)$  with  $r < R$ ,

$$G(z) \leq \frac{R+r}{R-r} G(w).$$

We apply Harnack's inequality to  $G(z) := 21n^{2c_2} - \log \alpha - \log |p|$  which is nonnegative harmonic on  $B(w, R)$  with  $R := r + \eta$ . By this inequality, we conclude that for all  $z \in B(w, r)$

$$(9.4) \quad 21n^{2c_2} - \log \alpha - \log |p(z)| \leq \frac{2r+\eta}{\eta} (21n^{2c_2} - \log \alpha - \log |p(w)|) \ll n^{2c_2} \frac{2r+\eta}{\eta}.$$

As  $\eta = .1n^{-1-c_2}$  and  $r < 1$ ,  $\frac{2r+\eta}{\eta} \ll \eta^{-1} \ll n^{c_2}$ . It follows that

$$\log |p(z)| + \log \alpha \gg -n^{3c_2}.$$

Together with (9.2), we have

$$|\log |p(z)| + \log \alpha| \ll n^{3c_2} \quad \forall z \in B(w, r).$$

Thus,

$$(9.5) \quad |\log |f(z)|| \ll n^{3c_2} \quad \forall z \in B(w, r).$$

By the triangle inequality and the definition of  $f$ ,

$$(9.6) \quad \|\log |F_n(z)|\|_{L^2(B(w,r))} \leq \|\log |f(z)|\|_{L^2(B(w,r))} + \sum_{i=1}^m \|\log |nz - n\zeta_i|\|_{L^2(B(w,r))}.$$

Notice that each of the  $m$  terms in the sum above is at most  $\int_{B(0, 2r-\eta)} |\log |nz||^2 dz$ , as  $|\zeta_i| \leq r - \eta$  for all  $i$ . As  $r < 1/n$ , we can further upper bound it by

$$\int_{B(0, 2/n)} |\log |nz||^2 dz = n^{-2} \int_{B(0, 2)} |\log |u||^2 du \ll n^{-2}.$$

Since  $m \leq 2n^{c_2}$ , we have

$$\|\log |F_n(z)|\|_{L^2(B(w,r))} \ll n^{-2+3c_2}$$

which implies the claim of the lemma as  $r \geq 1/(2n)$ .  $\square$

9.1. **Proof of Theorem 2.2.** With our  $\mathbf{P}_n^*$  plays the role of  $F_n$  in [31], the proof is identical to the proof of [31, Theorem 2.6] except that here we have a more relaxed condition C4 compared to [31, Condition C2(4)] in the second part on the growth of the second derivatives. In particular, in the latter, the condition was

$$\sum_{j=0}^n \sup_{z \in B_{\mathbb{C}}(s, 1/n)} |p_j''(a_n z)|^2 a_n^4 \leq C_1 n^{4+c_1} \sum_{j=0}^n |p_j(a_n s)|^2$$

while here we only require

$$\sup_{z \in B_{\mathbb{C}}(s, 1/n)} \sum_{j=0}^n |p_j''(a_n z)|^2 a_n^4 \leq C_1 n^{4+c_1} \sum_{j=0}^n |p_j(a_n s)|^2.$$

The only place in [31] where this condition was used is to prove [31, Equations 24,25]. Here, we show how to derive these equations with our new condition. We shall use the notation in [31] for this part which says

$$p(z) = \tilde{F}(z) - \tilde{F}(x) - (z-x)\tilde{F}'(x).$$

So,

$$\mathbb{E}p(z) = \mathbb{E}(\tilde{F}(z) - \tilde{F}(x) - (z-x)\tilde{F}'(z)) = 0.$$

This trivially implies that the left-most side of [31, Equation 24] is smaller than its right-most side.

Next, for [31, Equation 25], we observe that

$$p(z) = \int_x^z \tilde{F}'(t) dt - (z-x)\tilde{F}'(x) = \int_x^z \int_x^t \tilde{F}''(w) dw dt.$$

Hence, by Holder's inequality

$$|p(z)|^2 \leq \frac{(z-x)^2}{2} \int_x^z \int_x^t |\tilde{F}''(w)|^2 dw dt.$$

$$\text{Var}(p(z)) = \mathbb{E}|p(z)|^2 \leq \frac{(z-x)^4}{4} \sup_{w \in [x, z]} \mathbb{E}|\tilde{F}''(w)|^2 \leq \frac{(z-x)^4}{4} \sup_{w \in [x, z]} \sum_{j=1}^n |\phi_j''(w)|^2.$$

Now that we have moved the supremum outside of the summation, we can use our relaxed condition C4 to get that

$$\text{Var}(p(z)) = O\left(\delta_n^{4c_2 - c_1} \text{Var}(\tilde{F}_n(x))\right)$$

which is [31, Equation 25].

## 10. ALMOST SURE ZERO DISTRIBUTION RESULTS AND CONSEQUENCES

The main goal of this section is to prove Theorem 1.6. We start with several auxiliary facts on the asymptotic properties of random variables  $\{\xi_n\}_{n=0}^{\infty}$ .

**Lemma 10.1.** *If  $\{\xi_n\}_{n=0}^\infty$  are random variables such that  $\mathbb{E}[|\xi_n|^{2+\varepsilon_0}] < C$  for some constants  $C, \varepsilon_0 > 0$  and all  $n = 0, 1, \dots$ , then*

$$(10.1) \quad \limsup_{n \rightarrow \infty} |\xi_n|^{1/n} \leq 1 \quad a.s.$$

and

$$(10.2) \quad \limsup_{n \rightarrow \infty} \left( \max_{0 \leq k \leq n} |\xi_k| \right)^{1/n} \leq 1 \quad a.s.$$

*Proof of Lemma 10.1.* It is immediate from our assumptions that

$$x^{2+\varepsilon_0} \mathbb{P}(\{|\xi_n| > x\}) \leq \mathbb{E}[|\xi_n|^{2+\varepsilon_0}] < C.$$

Using the distribution function of  $|\xi_n|$  defined by  $H_n(x) = \mathbb{P}(\{|\xi_n| \leq x\})$ ,  $x \in \mathbb{R}$ , we estimate

$$1 - H_n(x) = \mathbb{P}(\{|\xi_n| > x\}) \leq C/x^{2+\varepsilon_0}, \quad x > 0, \quad n = 0, 1, \dots$$

Now (10.1) follows directly from Lemma 4.1 of [32], see (4.1) there.

Let  $\mathcal{A}$  be the event of probability one in (10.1). On  $\mathcal{A}$ , for any  $\varepsilon > 0$  there is  $n_\varepsilon \in \mathbb{N}$  such that  $|\xi_n|^{1/n} \leq 1 + \varepsilon$  for all  $n \geq n_\varepsilon$  by (10.1). Hence

$$\max_{0 \leq k \leq n} |\xi_k|^{1/n} \leq \max \left( \max_{0 \leq k \leq n_\varepsilon} |\xi_k|^{1/n}, 1 + \varepsilon \right) \rightarrow 1 + \varepsilon \quad \text{as } n \rightarrow \infty,$$

and (10.2) follows by letting  $\varepsilon \rightarrow 0$ . □

In the next two auxiliary lemmas, we also show that under our standard assumptions on the random variables, they cannot be too concentrated.

**Lemma 10.2.** *Let  $\{\xi_k\}_{k=0}^\infty$  be independent random variables such that  $\mathbb{E}[\xi_k] = 0$ ,  $\text{Var}[\xi_k] = 1$ , and  $\mathbb{E}[|\xi_k|^{2+\varepsilon_0}] < C$  for some constants  $C, \varepsilon_0 > 0$  and all  $k = 0, 1, \dots$ . There exists a constant  $a > 0$  depending only on  $C$  and  $\varepsilon_0$  such that for all  $k$ ,*

$$(10.3) \quad \mathbb{P}(|\xi_k| \leq a) \leq 1 - a.$$

Moreover, there exists a constant  $a' > 0$  depending only on  $C$  and  $\varepsilon_0$  such that for all  $k$  and for all  $x \in \mathbb{R}$ ,

$$(10.4) \quad \mathbb{P}(|\xi_k - x| \leq a') \leq 1 - a'.$$

*Proof.* To see (10.3), assume that  $\mathbb{P}(|\xi_k| \geq a) < a$ , then we obtain that

$$\begin{aligned} 1 &= \mathbb{E}[|\xi_k|^2] = \mathbb{E}[|\xi_k|^2 \mathbf{1}_{|\xi_k| < a}] + \mathbb{E}[|\xi_k|^2 \mathbf{1}_{|\xi_k| \geq a}] \\ &\leq a^2 + (\mathbb{E}[|\xi_k|^{2+\varepsilon_0}])^{2/(2+\varepsilon_0)} (\mathbb{P}(|\xi_k| \geq a))^{\varepsilon_0/(2+\varepsilon_0)} \quad \text{by Hölder's inequality} \\ &\leq a^2 + C^{2/(2+\varepsilon_0)} a^{\varepsilon_0/(2+\varepsilon_0)}. \end{aligned}$$

Since the latter upper bound converges to 0 as  $a \rightarrow 0$ , there must exist some  $a > 0$  (independent of  $k$ ) for which it is smaller than 1. This contradiction shows that there is a sufficiently small  $a > 0$  that satisfies (10.3).

Next, for (10.4), assume that  $\mathbb{P}(|\xi_k - x| \leq a') > 1 - a'$  then letting  $\xi'_k$  be an independent copy of  $\xi_k$ , we have

$$(10.5) \quad \mathbb{P}(|\xi_k - \xi'_k| \leq 2a') \geq (1 - a')^2 \geq 1 - 2a'.$$

By applying the first part to the random variables  $\frac{\xi_k - \xi'_k}{\sqrt{2}}$  which have mean 0, variance 1, and uniformly bounded  $(2 + \varepsilon_0)$ -moments, we obtain an  $a'$  that violates (10.5), and hence satisfies (10.4).  $\square$

**Lemma 10.3.** *If  $\{\xi_k\}_{k=0}^n$  are independent random variables such that  $\mathbb{E}[\xi_k] = 0$ ,  $\text{Var}[\xi_k] = 1$ , and  $\mathbb{E}[|\xi_k|^{2+\varepsilon_0}] < C$  for some constants  $C, \varepsilon_0 > 0$  and all  $k = 0, 1, \dots$ , then there is  $b > 0$  such that*

$$(10.6) \quad \liminf_{n \rightarrow \infty} \left( \max_{n-b \log n < k \leq n} |\xi_k| \right)^{1/n} \geq 1 \quad a.s.$$

*Proof.* Let  $s_n \leq n$ ,  $n \in \mathbb{N}$ , be a sequence of natural numbers that will be specified later. Consider

$$M_n := \max_{n-s_n < k \leq n} |\xi_k|.$$

The statement

$$\liminf_{n \rightarrow \infty} (M_n)^{1/n} \geq 1 \quad a.s.$$

is equivalent to

$$\mathbb{P}(\{M_n \leq \lambda^n \text{ i.o.}\}) = 0$$

for all positive  $\lambda < 1$ . The latter would follow from the first Borel-Cantelli Lemma, if we show that

$$\sum_{n=1}^{\infty} \mathbb{P}(\{M_n \leq \lambda^n\}) < \infty$$

for all positive  $\lambda < 1$ . Since the variables  $\{\xi_k\}_{k=0}^{\infty}$  are independent, we have

$$\mathbb{P}(\{M_n \leq \lambda^n\}) = \prod_{k=1}^{s_n} \mathbb{P}(\{|\xi_{n-k}| \leq \lambda^n\}).$$

Using (10.3), for any  $\lambda < 1$  we find  $N = N(a, \lambda) \in \mathbb{N}$  such that  $\mathbb{P}(\{|\xi_k| \leq \lambda^n\}) \leq 1 - a$  for all  $n \geq N$  and all  $k = 0, 1, \dots$ . This gives

$$\sum_{n=1}^{\infty} \mathbb{P}(\{M_n \leq \lambda^n\}) \leq \sum_{n=1}^{\infty} (1 - a)^{s_n} < \infty,$$

provided  $(1 - a)^{s_n} \leq 1/n^2$  for large  $n$ . It suffices to take  $s_n \geq (-2/\log(1 - a)) \log n$  to satisfy the latter condition.  $\square$

We now state a companion result for the coefficients of our contracted random orthogonal polynomials

$$P_n^*(x) = P_n(a_n x) = \sum_{k=0}^n \xi_k p_k(a_n x) = \sum_{k=0}^n c_{k,n} x^k,$$

see (1.4).

**Lemma 10.4.** *Let  $\{\xi_k\}_{k=0}^n$  be independent random variables such that  $\mathbb{E}[\xi_k] = 0$ ,  $\text{Var}[\xi_k] = 1$ , and  $\mathbb{E}[|\xi_k|^{2+\varepsilon_0}] < C$  for some constants  $C, \varepsilon_0 > 0$  and all  $k = 0, 1, \dots$ . Assume that  $W = e^{-Q} \in \mathcal{F}(C^2)$ , where  $Q$  is even, and the function  $T$  in the definition of  $\mathcal{F}(C^2)$  satisfies (1.1) with  $\alpha \in (1, \infty)$ . For the orthogonal polynomials  $p_k$  that correspond to the weight  $W$ , write*

$$p_k(a_n x) = \sum_{j=0}^k b_{j,k} x^j = b_{k,k} \prod_{j=1}^k (x - x_{j,k}), \quad k = 0, 1, \dots, n.$$

If the random variables  $c_{k,n}$  are defined by

$$c_{k,n} := \sum_{i=k}^n \xi_i b_{k,i}, \quad k = 0, 1, \dots, n,$$

then there is  $b > 0$  such that

$$(10.7) \quad \liminf_{n \rightarrow \infty} \left( \max_{n-b \log n < k \leq n} |c_{k,n}| \right)^{1/n} \geq 2 e^{1/\alpha} \quad a.s.$$

*Proof.* We follow the idea in [7, Lemma 4.2]. Note that  $b_{k,k} = a_n^k \gamma_k$ , where  $\gamma_k$  is the leading coefficient of  $p_k$ . We have

$$(10.8) \quad \lim_{n \rightarrow \infty} |b_{n,n}|^{1/n} = \lim_{n \rightarrow \infty} a_n \gamma_n^{1/n} = 2 e^{1/\alpha}.$$

For every  $\varepsilon > 0$ , it suffices to show that

$$(10.9) \quad \mathbb{P} \left( \left( \max_{n-b \log n < k \leq n} |c_{k,n}| \right)^{1/n} \leq 2 e^{1/\alpha} - \varepsilon \text{ infinitely often} \right) = 0.$$

Note that  $|c_{k,n}| \leq (2e^{1/\alpha} - \varepsilon)^n$  implies that  $\xi_k b_{k,k}$  lies in an interval of radius  $(2e^{1/\alpha} - \varepsilon)^n$  centered around  $\sum_{i=k+1}^n \xi_i b_{k,i}$ . Thus, by independence, for any event  $A_{k,n} \in \sigma(\xi_{k+1}, \dots, \xi_n)$  – the  $\sigma$ -algebra generated by  $\xi_{k+1}, \dots, \xi_n$ , we have

$$\mathbb{P}(|c_{k,n}|^{1/n} \leq 2 e^{1/\alpha} - \varepsilon | A_{k,n}) \leq \sup_{x \in \mathbb{R}} \mathbb{P}(|\xi_k b_{k,k} - x| \leq (2 e^{1/\alpha} - \varepsilon)^n).$$

By (10.8), this implies that for sufficiently large  $n$  and  $k \geq n - b \log n$ ,

$$\mathbb{P}(|c_{k,n}|^{1/n} \leq 2 e^{1/\alpha} - \varepsilon | A_{k,n}) \leq \sup_{x \in \mathbb{R}} \mathbb{P}(|\xi_k - x| \leq (1 - \delta)^n)$$

where  $\delta = \frac{\varepsilon}{2 e^{1/\alpha}}$  is a constant.

By (10.4), the last quantity is smaller than  $1 - a'$  for the same constant  $a'$  as in (10.4), again for all sufficiently large  $n$ . Thus, by setting

$$A_{k,n} = \{|c_{\ell,n}|^{1/n} \leq 2 e^{1/\alpha} - \varepsilon, \forall \ell \in (k, n] \cap \mathbb{N}\} \in \sigma(\xi_{k+1}, \dots, \xi_n)$$

and  $A_{n,n}$  be the whole sample space, we get

$$\begin{aligned} \mathbb{P} \left( \left( \max_{n-b \log n < k \leq n} |c_{k,n}| \right)^{1/n} \leq 2 e^{1/\alpha} - \varepsilon \right) &= \prod_{k=[n-b \log n]+1}^n \mathbb{P}(|c_{k,n}| \leq (2 e^{1/\alpha} - \varepsilon)^n | A_{k,n}) \\ &\leq (1 - a')^{b \log n} \leq n^{-a'b}. \end{aligned}$$

By choosing  $b = 2/a'$ , the right-most side becomes  $n^{-2}$  which is summable over  $n$ . By the first Borel-Cantelli Lemma, this implies (10.9).  $\square$

*Proof of Theorem 1.6.* This proof is similar to the proof of Theorem 2.4 in [28], but it has several additional steps due to different assumptions on random coefficients. Recall that the standard Freud weight with index  $\alpha$  is given by

$$w(s) = e^{-\gamma_\alpha |s|^\alpha}, \quad s \in \mathbb{R},$$

where

$$\gamma_\alpha = \frac{\Gamma(\frac{\alpha}{2})\Gamma(\frac{1}{2})}{2\Gamma(\frac{\alpha}{2} + \frac{1}{2})} = \int_0^1 \frac{t^{\alpha-1}}{\sqrt{1-t^2}} dt$$

see [35, p. 239]. Note that by [35, p. 240],  $F_w = \log 2 + 1/\alpha$  is the modified Robin constant and  $\mu_w = \mu_\alpha$  is the equilibrium measure corresponding to  $w$ . Following [35], we call a sequence of monic polynomials  $\{M_n\}_{n=1}^\infty$ , with  $\deg(M_n) = n$ , asymptotically extremal with respect to the weight  $w$  if it satisfies

$$(10.10) \quad \lim_{n \rightarrow \infty} \|w^n M_n\|_{L^\infty(\mathbb{R})}^{1/n} = e^{-F_w} = e^{-1/\alpha}/2.$$

Theorem 4.2 of [35, p. 170] states that the normalized zero counting measures of  $M_n$  converge weak\* to  $\mu_w = \mu_\alpha$ . On the other hand, by Corollary 2.6 of [35, p. 157] and Theorem 5.1 of [35, p. 240],

$$\|w^n M_n\|_{L^\infty(\mathbb{R})} = \|w^n M_n\|_{L^\infty([-1,1])}.$$

Applying Theorem 3.6 of [35, p. 46], we see that (10.10) is equivalent to

$$(10.11) \quad \limsup_{n \rightarrow \infty} \|w^n M_n\|_{L^\infty([-1,1])}^{1/n} \leq e^{-F_w} = e^{-1/\alpha}/2.$$

We construct the monic polynomials  $M_n$  from  $P_n^*$  in such a way that they share most of zeros, and therefore have the same limit of their zero counting measures. Continuing with our previous notation, we write

$$P_n^*(z) = P_n(a_n z) = \sum_{k=0}^n \xi_k p_k(a_n z) = \sum_{k=0}^n c_{k,n} z^k = c_{d_n,n} \prod_{k=1}^{d_n} (z - z_{k,n}), \quad c_{d_n,n} \neq 0,$$

where we assume that the zeros  $\{z_{k,n}\}_{k=1}^{d_n}$  are listed in the order of decreasing absolute values:  $|z_{1,n}| \geq |z_{2,n}| \geq \dots \geq |z_{d_n,n}|$ , and where  $d_n$  is the actual degree of  $P_n^*$ . Note that  $n - d_n = o(n)$  with probability one by (10.7). Let  $m_n$  be the number of zeros of  $P_n^*$  with absolute value greater than 2, and consider  $l_n = \min(m_n, b \log n)$ , where  $b > 0$  is a constant from Lemma 10.4. We define the monic polynomial  $M_n$  of degree  $n$  by

$$M_n(z) := \frac{(z-2)^{n-d_n+l_n} P_n^*(z)}{c_{d_n,n} \prod_{1 \leq k \leq l_n} (z - z_{k,n})}, \quad n \in \mathbb{N},$$

and show they are asymptotically extremal in the sense of (10.10)-(10.11) with probability one. Thus the zero counting measures of  $M_n$  converge weak\* to  $\mu_\alpha$  as discussed above, and the same conclusion is true for the zero counting measures of  $P_n^*$ , because the sets of roots for  $M_n$  and  $P_n^*$  differ only by  $o(n)$  elements.

Using Vieta's formulas for  $P_n^*$ , we obtain that

$$|c_{n-k,n}| \leq |c_{d_n,n}| \left| \sum_{j_1 < \dots < j_k} z_{j_1,n} \cdots z_{j_k,n} \right| \leq |c_{d_n,n}| \binom{n}{k} \prod_{j=1}^k |z_{j,n}| \leq n^k |c_{d_n,n}| \prod_{j=1}^k |z_{j,n}|, \quad k \leq l_n,$$

where we used decreasing ordering of zeros. Since  $l_n = O(\log n)$ , it follows from the last inequality and Lemma 10.4 that

$$\liminf_{n \rightarrow \infty} \left( |c_{d_n,n}| \prod_{j=1}^k |z_{j,n}| \right)^{1/n} \geq \liminf_{n \rightarrow \infty} \left( \max_{1 \leq k \leq l_n} |c_{n-k,n}| \right)^{1/n} \geq 2e^{1/\alpha} \quad \text{a.s.}$$

Further, we apply the above inequality and that  $n - d_n + l_n = o(n)$  to estimate (with probability one)

$$\begin{aligned} \limsup_{n \rightarrow \infty} \|M_n w^n\|_{L^\infty([-1,1])}^{1/n} &\leq \limsup_{n \rightarrow \infty} \|P_n^* w^n\|_{L^\infty([-1,1])}^{1/n} \left\| \frac{(z-2)^{n-d_n+l_n}}{c_{d_n,n} \prod_{1 \leq k \leq l_n} (z - z_{k,n})} \right\|_{L^\infty([-1,1])}^{1/n} \\ &\leq \limsup_{n \rightarrow \infty} \|P_n^* w^n\|_{L^\infty([-1,1])}^{1/n} \left( \frac{3^{n-d_n+l_n}}{|c_{d_n,n}| \prod_{1 \leq k \leq l_n} (|z_{k,n}|/2)} \right)^{1/n} \\ &\leq \frac{1}{2e^{1/\alpha}} \limsup_{n \rightarrow \infty} \|P_n^* w^n\|_{L^\infty([-1,1])}^{1/n} \\ &= e^{-F_w} \limsup_{n \rightarrow \infty} \|P_n^* w^n\|_{L^\infty([-1,1])}^{1/n}. \end{aligned}$$

For the second inequality in the above display, we also used that  $|z_{k,n}| \geq 2$ ,  $1 \leq k \leq l_n$ , by our construction, so that  $|z - z_{k,n}| \geq |z_{k,n}| - 1 \geq |z_{k,n}|/2$  for  $z \in [-1, 1]$ ,  $1 \leq k \leq l_n$ , and clearly  $|z - 2| \leq 3$  for  $z \in [-1, 1]$ .

Thus the proof would be completed if we show that

$$(10.12) \quad \limsup_{n \rightarrow \infty} \|P_n^* w^n\|_{L^\infty([-1,1])}^{1/n} \leq 1 \quad \text{a.s.}$$

Recall that by (4.1)

$$\|W P_n\|_{L^\infty(\mathbb{R})} \leq Cn \left( \sum_{k=0}^n |\xi_k|^2 \right)^{1/2}, \quad n \in \mathbb{N}.$$

Changing variable  $x = a_n s$  and introducing  $w_n(s) = W(a_n s)^{1/n}$ , we obtain that

$$\|P_n^* w_n^n\|_{L^\infty([-1,1])} = \|P_n W\|_{L^\infty([-a_n, a_n])} = \|P_n W\|_{L^\infty(\mathbb{R})}$$

by [22, p. 4]. Hence

$$\limsup_{n \rightarrow \infty} \|P_n^* w_n^n\|_{L^\infty([-1,1])}^{1/n} \leq \limsup_{n \rightarrow \infty} \left( \max_{0 \leq k \leq n} |\xi_k| \right)^{1/n} \leq 1 \quad \text{a.s.}$$

by (10.2) of Lemma 10.1. It follows that, with probability one, we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \|P_n^* w^n\|_{L^\infty([-1,1])}^{1/n} &\leq \limsup_{n \rightarrow \infty} \|P_n^* w_n^n\|_{L^\infty([-1,1])}^{1/n} \|w/w_n\|_{L^\infty([-1,1])} \\ &\leq \limsup_{n \rightarrow \infty} \|w/w_n\|_{L^\infty([-1,1])}. \end{aligned}$$

Since  $w_n$  and  $w$  are both even, to prove (10.12) it only remains to show that

$$\limsup_{n \rightarrow \infty} \|w/w_n\|_{L^\infty([0,1])} \leq 1.$$

But exactly this limit relation was established in the end of proof for Theorem 2.4 in [28], so that we refer there to avoid duplication.  $\square$

*Proof of Corollary 1.7.* The proof of this result is identical to the proof of Corollary 2.5 in [28]. Thus we omit it.  $\square$

## 11. PROOF OF THEOREM 2.4

We recall that  $D(\varepsilon) = [-1 + \varepsilon, 1 - \varepsilon]$ . We need to show that the number of real roots outside of  $D(\varepsilon)$  is negligible, i.e.,

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N_{P_n^*}(\mathbb{R} \setminus D(\varepsilon))]}{n} = 0.$$

The corresponding result for  $\mathbb{E}[N_{\tilde{P}_n^*}]$  is included in the above statement.

For any  $\varepsilon > 0$ , let  $F_\varepsilon := ([-1/\varepsilon, \varepsilon - 1] \cup [1 - \varepsilon, 1/\varepsilon]) \times [-\varepsilon, \varepsilon]$ . Recall from (1.3) that  $\mu_\alpha$  is absolutely continuous, and  $\text{supp } \mu_\alpha = [-1, 1]$ . It follows that  $\mu_\alpha(\partial F_\varepsilon) = 0$ , and we can apply Corollary 1.7 to obtain that

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}N_{P_n^*}(F_\varepsilon)}{n} = \mu_\alpha(F_\varepsilon).$$

Furthermore, monotone convergence and basic measure properties give that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N_{P_n^*}(\mathbb{R} \setminus D(\varepsilon))]}{n} &\leq \limsup_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N_{P_n^*}(F_\varepsilon)]}{n} = \limsup_{\varepsilon \rightarrow 0} \mu_\alpha(F_\varepsilon) \\ &= \limsup_{\varepsilon \rightarrow 0} \mu_\alpha([-1, \varepsilon - 1] \cup [1 - \varepsilon, 1]) = 0, \end{aligned}$$

which completes this proof.

## 12. CORRELATIONS OF REAL ROOTS

For each  $1 \leq k \leq n$ , let  $\rho_k(x_1, \dots, x_k)$  be the  $k$ -correlation of the real roots of  $P_n(x)$  (see for instance [16]), for which

$$\mathbb{E} \left[ \sum G(\zeta_{i_1}, \dots, \zeta_{i_k}) \right] = \int_{\mathbb{R}^k} G(x_1, \dots, x_k) \rho_k(x_1, \dots, x_k) dx_1 \dots dx_k,$$

for any continuous, compactly supported test function  $G : \mathbb{R}^k \rightarrow \mathbb{R}$ , where the sum runs over all  $k$ -tuples  $(\zeta_{i_1}, \dots, \zeta_{i_k})$  of the real roots of  $P_n(x)$ .

In principle one can use Kac-Rice formula (see also [16]<sup>1</sup>) to compute these correlation functions, that

$$(12.1) \quad \rho_k(x_1, \dots, x_k) = \int_{\mathbb{R}^k} |y_1, \dots, y_k| p(\mathbf{0}, \mathbf{y}) dy_1 \dots dy_k,$$

where  $p(\cdot)$  is the joint density function of the random vectors  $(P_n(x_1), \dots, P_n(x_k))$  and  $(P'_n(x_1), \dots, P'_n(x_k))$ .

This formula is especially useful when  $\xi_i$  are iid standard gaussian and when  $k$  is small. For instance, directly related to our current setting of orthogonal polynomials, the cases  $k = 1$  and  $k = 2$  were computed in [28] and [26] respectively. However, the formulas become increasingly more complicated when  $k$  gets larger and for other ensembles of random coefficients. Our goal in this section is to provide some alternative formulas by following [15].

For each  $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{R}^k$ , denote by  $V(\mathbf{x})$  the Vandermonde-type matrix

$$V(\mathbf{x}) = \begin{pmatrix} p_0(x_1) & p_1(x_1) & \dots & p_{k-1}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_0(x_k) & p_1(x_k) & \dots & p_{k-1}(x_k) \end{pmatrix},$$

where we recall that  $p_i(x)$  are the orthogonal polynomials of degree  $i$  with respect to  $\mu$ . Assume that  $p_i(x) = \gamma_i x^i + \dots$ , then from the recurrence formula (7.1) we have  $\gamma_i = (A_0 \dots A_{i-1})^{-1} \gamma_0$  and

$$\det(V(\mathbf{x})) = \prod_{m=0}^{k-1} \gamma_m \prod_{1 \leq i < j \leq k} (x_j - x_i).$$

As the random coefficients  $\xi_i$  of  $P_n(x)$  have density, with probability one the roots of  $P_n(x)$  are distinct, and hence we will assume that the  $x_i$  are distinct. Consider the random function  $\boldsymbol{\eta} = (\eta_0, \dots, \eta_{k-1})^T : \mathbb{R}^k \rightarrow \mathbb{R}^k$  defined as

$$(12.2) \quad \boldsymbol{\eta}(\mathbf{x}) := -V^{-1}(\mathbf{x}) \begin{pmatrix} \sum_{j=k}^n \xi_j p_j(x_1) \\ \vdots \\ \sum_{j=k}^n \xi_j p_j(x_k) \end{pmatrix}.$$

Our main result of this section is the following formula.

**Theorem 12.1.** *Let  $f_0, \dots, f_{k-1}$  be the density functions of  $\xi_0, \dots, \xi_{k-1}$  respectively. With  $\boldsymbol{\eta}$  from (12.2) we have*

$$\rho_k(\mathbf{x}) = \prod_{m=0}^{k-1} \gamma_m^{-1} \prod_{1 \leq i < j \leq k} |x_i - x_j|^{-1} \mathbb{E} \left( \prod_{i=1}^k \left| \sum_{j=0}^{k-1} \eta_j(\mathbf{x}) p_j(x_i) + \sum_{j=k}^n \xi_j p_j(x_i) \right| \prod_{i=0}^{k-1} f_i(\eta_i(\mathbf{x})) \right),$$

where the expectation is with respect to  $\xi_k, \dots, \xi_n$ .

Alternatively, we also have following analog of [15, Theorem 2.3], where  $\sigma_i(\mathbf{x}), 1 \leq i \leq k$ , are the elementary symmetric polynomials  $\sum_{1 \leq j_1 < \dots < j_i \leq k} x_{j_1} x_{j_2} \dots x_{j_i}$  (with the convention that

<sup>1</sup>The literature on this formula is so vast that it is impossible to list even a small portion of it.

$\sigma_0(\mathbf{x}) = 1$ ) and where for short we write

$$\langle p(x), q(x) \rangle_\mu := \int_{\mathbb{R}} p(x)q(x)d\mu(x).$$

**Theorem 12.2.** *Let  $f_0, \dots, f_n$  be the density functions of  $\xi_0, \dots, \xi_n$  respectively. We have*

$$\begin{aligned} \rho_k(\mathbf{x}) &= \prod_{m=0}^n \gamma_m^{-1} \prod_{1 \leq i < j \leq k} |x_i - x_j| \times \\ &\times \int_{\mathbb{R}^{n-k+1}} \prod_{l=0}^n f_l \left( \sum_{i=l}^n \left( \sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(\mathbf{x}) t_j \right) \langle x^i, p_l(x) \rangle_\mu \right) \prod_{i=1}^k \left| \sum_{j=0}^{n-k} t_j x_i^j \right| dt_0 \dots dt_{n-k}, \end{aligned}$$

These formulas seem to be useful when  $k$  is comparable to  $n$ . For instance when  $k = n$  we obtain the following joint density formula for  $n$  real roots of  $P_n(x)$ .

**Corollary 12.3.**

$$\rho_n(\mathbf{x}) = \prod_{m=0}^n \gamma_m^{-1} \prod_{1 \leq i < j \leq n} |x_i - x_j| \int_{\mathbb{R}} \prod_{l=0}^n f_l \left( \left( \sum_{i=l}^n (-1)^{n-i} \sigma_{n-i}(\mathbf{x}) \langle x^i, p_l(x) \rangle_\mu \right) t \right) t^n dt.$$

In what follows we discuss the proof of the above results. Notice that  $P_n(x_1) = \dots = P_n(x_k) = 0$  if and only if

$$\begin{pmatrix} p_0(x_1) & p_1(x_1) & \dots & p_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_0(x_k) & p_1(x_k) & \dots & p_n(x_k) \end{pmatrix} \begin{pmatrix} \xi_0 \\ \vdots \\ \xi_n \end{pmatrix} = \mathbf{0},$$

which can be rewritten as

$$\begin{pmatrix} p_k(x_1) & p_{k+1}(x_1) & \dots & p_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_k(x_k) & p_{k+1}(x_k) & \dots & p_n(x_k) \end{pmatrix} \begin{pmatrix} \xi_k \\ \vdots \\ \xi_n \end{pmatrix} = -V(\mathbf{x}) \begin{pmatrix} \xi_0 \\ \vdots \\ \xi_{k-1} \end{pmatrix}.$$

Thus

$$\boldsymbol{\eta}(\mathbf{x}) = \begin{pmatrix} \xi_0 \\ \vdots \\ \xi_{k-1} \end{pmatrix}.$$

Denote by  $J_{\boldsymbol{\eta}}(\mathbf{x})$  the Jacobian matrix of  $\boldsymbol{\eta}$  at the point  $\mathbf{x}$ .

**Lemma 12.4.** *We have*

$$\det(J_{\boldsymbol{\eta}}(\mathbf{x})) = (-1)^k \frac{\prod_{i=1}^k \left( \sum_{j=0}^{k-1} \eta_j(\mathbf{x}) p'_j(x_i) + \sum_{j=k}^n \xi_j p'_j(x_i) \right)}{\prod_{m=0}^{k-1} \gamma_m \prod_{1 \leq i < j \leq k} (x_j - x_i)}.$$

*Proof.* Recall that

$$V(\mathbf{x})\boldsymbol{\eta}(\mathbf{x}) = - \begin{pmatrix} p_k(x_1) & p_{k+1}(x_1) & \cdots & p_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_k(x_k) & p_{k+1}(x_k) & \cdots & p_n(x_k) \end{pmatrix} \begin{pmatrix} \xi_k \\ \vdots \\ \xi_n \end{pmatrix}.$$

By differentiating,

$$\begin{aligned} V(\mathbf{x})J_{\boldsymbol{\eta}}(\mathbf{x}) &+ \text{diag} \left( \sum_{j=0}^{k-1} \eta_j(\mathbf{x})p'_j(x_1), \dots, \sum_{j=0}^{k-1} \eta_j(\mathbf{x})p'_j(x_k) \right) \\ &= - \text{diag} \left( \sum_{j=k}^n \xi_j p'_j(x_1), \dots, \sum_{j=k}^n \xi_j p'_j(x_k) \right). \end{aligned}$$

It thus follows that

$$\det(J_{\boldsymbol{\eta}}(\mathbf{x})) = (-1)^k \det(V(\mathbf{x}))^{-1} \prod_{i=1}^k \left( \sum_{j=0}^{k-1} \eta_j(\mathbf{x})p'_j(x_i) + \sum_{j=k}^n \xi_j p'_j(x_i) \right).$$

□

We next rely on the Coarea formula [12].

**Lemma 12.5.** *Let  $B \subset \mathbb{R}^k$  be a region. Let  $u : B \rightarrow \mathbb{R}^k$  be a Lipschitz function and  $h : \mathbb{R}^k \rightarrow \mathbb{R}^1$  be an  $L^1$ -function. Then*

$$\int_{\mathbb{R}^k} \#\{\mathbf{x} \in B : u(\mathbf{x}) = \mathbf{y}\} h(\mathbf{y}) d\mathbf{y} = \int_B |\det J_u(\mathbf{x})| h(u(\mathbf{x})) d\mathbf{x},$$

where  $J_u(\mathbf{x})$  is the Jacobian matrix of  $u(\mathbf{x})$ .

*Proof of Theorem 12.1.* Let  $B_1, \dots, B_k$  be a family of mutually disjoint Borel subsets in  $\mathbb{R}$  and let  $B = B_1 \times \cdots \times B_k$ . By the Coarea formula and by Fubini theorem we have

$$\begin{aligned} \mathbb{E} \left[ \prod_{i=1}^k N_n(B_i) \right] &= \mathbb{E} \#\{\mathbf{x} \in B : \boldsymbol{\eta}(\mathbf{x}) = (\xi_0, \dots, \xi_{k-1})\} \\ &= \mathbb{E} \int_{\mathbb{R}^k} \#\{\mathbf{x} \in B : \boldsymbol{\eta}(\mathbf{x}) = \mathbf{y}\} f_0(y_0) \cdots f_{k-1}(y_{k-1}) d\mathbf{y} \\ &= \int_B \mathbb{E} |\det(J_{\boldsymbol{\eta}}(\mathbf{x}))| f_0(\eta_0(\mathbf{x})) \cdots f_{k-1}(\eta_{k-1}(\mathbf{x})) d\mathbf{x}, \end{aligned}$$

where we recall that  $f_0, \dots, f_{k-1}$  are the density functions of  $\xi_0, \dots, \xi_{k-1}$  respectively. □

*Proof of Theorem 12.2.* Note that by Theorem 12.1,

$$(12.3) \quad \begin{aligned} \rho_k(\mathbf{x}) &= \prod_{m=0}^{k-1} \gamma_m^{-1} \prod_{1 \leq i < j \leq k} |x_j - x_i|^{-1} \\ &\times \int_{\mathbb{R}^{n-k+1}} \prod_{i=1}^k \left| \sum_{j=0}^n a_j p'_j(x_i) \right| \prod_{i=0}^n f_i(a_i) da_k da_{k+1} \dots da_n, \end{aligned}$$

where  $(a_0, \dots, a_{k-1})^T = -V^{-1}(\mathbf{x})(\sum_{j=k}^n a_j p_j(x_1), \dots, \sum_{j=k}^n a_j p_j(x_k))^T$ .

This means that  $x_1, x_2, \dots, x_k$  are zeros of the polynomial  $\sum_{i=0}^n a_i p_i(x)$ . Hence there exists a unique polynomial  $\sum_{j=0}^{n-k} b_j x^j$  such that

$$\begin{aligned} P(x) &= \sum_{i=0}^n a_i p_i(x) = \prod_{i=1}^k (x - x_i) \left( \sum_{j=0}^{n-k} b_j x^j \right) = \left( \sum_{j=0}^k (-1)^{k-j} \sigma_{k-j}(\mathbf{x}) x^j \right) \left( \sum_{j=0}^{n-k} b_j x^j \right) \\ &= \sum_{i=0}^n \left( \sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(\mathbf{x}) b_j \right) x^i. \end{aligned}$$

The variables  $a_0, \dots, a_n$  are uniquely defined by  $\mathbf{x}$  and  $b_0, \dots, b_{n-k}$  from the above equation, that for any  $0 \leq l \leq n$  we have

$$a_l = \langle P(x), p_l(x) \rangle_\mu = \sum_{i=l}^n \left( \sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(\mathbf{x}) b_j \right) \langle x^i, p_l(x) \rangle_\mu$$

where we note that  $\langle x^i, p_l(x) \rangle_\mu = 0$  if  $i \leq l - 1$ . In particular,

$$\begin{aligned} a_k &= \sum_{i=k}^n \left( \sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(\mathbf{x}) b_j \right) \langle x^i, p_k(x) \rangle_\mu \\ &= \left( \sum_{j=0}^{n-k} (-1)^j \sigma_j(\mathbf{x}) b_j \right) \langle x^k, p_k(x) \rangle_\mu + \left( \sum_{j=1}^{n-k} (-1)^{j-1} \sigma_{j-1}(\mathbf{x}) b_j \right) \langle x^{k+1}, p_k(x) \rangle_\mu + \dots \\ &\quad + \left( \sum_{j=n-k}^{n-k} (-1)^{k-n+j} \sigma_{k-n+j}(\mathbf{x}) b_j \right) \langle x^n, p_k(x) \rangle_\mu \\ &= b_0 \langle x^k, p_k(x) \rangle_\mu + b_1 c_{k,1} + \dots + b_{n-k} c_{k,n-k}, \end{aligned}$$

for some numbers  $c_{k,1}, \dots, c_{k,n-k}$  independently of the  $b_i$ .

More generally, for any  $0 \leq l \leq n - k$

$$\begin{aligned}
a_{k+l} &= \sum_{i=k+l}^n \left( \sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(\mathbf{x}) b_j \right) \langle x^i, p_{k+l}(x) \rangle \\
&= \left( \sum_{j=l}^{n-k} (-1)^{j-l} \sigma_{j-l}(\mathbf{x}) b_j \right) \langle x^{k+l}, p_{k+l}(x) \rangle \\
&\quad + \left( \sum_{j=l+1}^{n-k} (-1)^{j-l-1} \sigma_{j-l-1}(\mathbf{x}) b_j \right) \langle x^{k+l+1}, p_{k+l}(x) \rangle + \cdots + \\
&\quad + \left( \sum_{j=n-k}^{n-k} (-1)^{k-n+j} \sigma_{k-n+j}(\mathbf{x}) b_j \right) \langle x^n, p_{k+l}(x) \rangle \\
&= b_l \langle x^{k+l}, p_{k+l}(x) \rangle + b_{l+1} c_{k+l,1} + \cdots + b_{n-k} c_{k+l,n-k-l}
\end{aligned}$$

for some numbers  $c_{k+l,1}, \dots, c_{k,n-k-l}$  independently of the  $b_i$ .

Thus the Jacobian of the substitution of  $(a_k, \dots, a_n)$  by  $(b_0, \dots, b_{n-k})$  is a lower triangle matrix with diagonal  $\text{diag}(\langle x^k, p_k(x) \rangle_\mu, \dots, \langle x^n, p_n(x) \rangle_\mu) = \text{diag}(\gamma_k^{-1}, \dots, \gamma_n^{-1})$ . Hence

$$(12.4) \quad da_k \dots da_n = \prod_{j=k}^n \gamma_j^{-1} db_0 \dots db_{n-k}.$$

Finally, by differentiating  $P(x) = \sum_{j=0}^n a_j p_j(x) = \prod_{j=1}^k (x - x_j) \left( \sum_{j=0}^{n-k} b_j x^j \right)$  at the point  $x_i$  we get

$$\sum_{j=0}^n a_j p'_j(x_i) = \prod_{j \in \{1, \dots, k\} \setminus \{i\}} (x_j - x_i) \left( \sum_{j=0}^{n-k} b_j x_i^j \right), \quad i = 1, \dots, k.$$

Substitute the above and (12.4) into (12.3) we obtain the claim.  $\square$

### Acknowledgements.

This paper was written within the AIM SQuaREs project ‘‘Random Polynomials.’’ All authors would like to gratefully acknowledge support of the American Institute of Mathematics. Y. Do was supported by DMS-1800855. D. Lubinsky was supported by NSF grant DMS-1800251. H. H. Nguyen was supported by NSF CAREER grant DMS-1752345. O. Nguyen was supported by NSF grant DMS-2125031. I. Pritsker was partially supported by NSA grant H98230-21-1-0008, NSF grant DMS-2152935, and by the Vaughn Foundation endowed Professorship in Number Theory.

### REFERENCES

- [1] Jürgen Angst and Guillaume Poly. Universality of the mean number of real zeros of random trigonometric polynomials under a weak Cramer condition. *arXiv preprint arXiv:1511.08750*, 2015.
- [2] Jean-Marc Azaïs, Federico Dalmao, José León, Ivan Nourdin, and Guillaume Poly. Local universality of the number of zeros of random trigonometric polynomials with continuous coefficients. *arXiv preprint arXiv:1512.05583*, 2015.

- [3] Vlad Bally, Lucia Caramellino, and Guillaume Poly. Non universality for the variance of the number of real roots of random trigonometric polynomials. *Probability Theory and Related Fields*, 174(3):887–927, 2019.
- [4] André Bloch and György Pólya. On the roots of certain algebraic equations. *Proceedings of the London Mathematical Society*, 2(1):102–114, 1932.
- [5] Minaketan Das. Real zeros of a random sum of orthogonal polynomials. *Proceedings of the American Mathematical Society*, 27(1):147–153, 1971.
- [6] Minaketan Das and SS Bhatt. Real roots of random harmonic equations. *Indian J. Pure Appl. Math*, 13:411–420, 1982.
- [7] Duncan Dauvergne. A necessary and sufficient condition for convergence of the zeros of random polynomials. *Advances in Mathematics*, 384:107691, 2021.
- [8] Yen Do, Hoi Nguyen, and Van Vu. Real roots of random polynomials: expectation and repulsion. *Proceedings of the London Mathematical Society*, 111(6):1231–1260, 2015.
- [9] Yen Do, Oanh Nguyen, and Van Vu. Roots of random polynomials with coefficients with polynomial growth. *Annals of Probability*, 46(5):2407–2494, 2018.
- [10] Yen Do, Oanh Nguyen, and Van Vu. Random orthonormal polynomials: local universality and expected number of real roots. *arXiv preprint arXiv:2012.10850*, 2020.
- [11] Paul Erdős and Albert Offord. On the number of real roots of a random algebraic equation. *Proceedings of the London Mathematical Society*, 3(1):139–160, 1956.
- [12] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York, Inc., New York, 1969.
- [13] Hendrik Flasche. Expected number of real roots of random trigonometric polynomials. *Stochastic Processes and their Applications*, 2017.
- [14] Géza Freud. *Orthogonal Polynomials*. 1971.
- [15] Friedrich Götze, Dzianis Kaliada, and Dmitry Zaporozhets. Correlation functions of real zeros of random polynomials. *arXiv preprint arXiv:1510.00025*, 2015.
- [16] John Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51. American Mathematical Society Providence, RI, 2009.
- [17] Alexander Iksanov, Zakhar Kabluchko, and Alexander Marynych. Local universality for real roots of random trigonometric polynomials. *J. London Math. Soc.*, 21, 2016.
- [18] Zakhar Kabluchko and Dmitry Zaporozhets. Asymptotic distribution of complex zeros of random analytic functions. *The Annals of Probability*, 42(4):1374–1395, 2014.
- [19] Mark Kac. On the average number of real roots of a random algebraic equation. *Bulletin of the American Mathematical Society*, 49(1):314–320, 1943.
- [20] Eli Levin and Doron Lubinsky. Applications of universality limits to zeros and reproducing kernels of orthogonal polynomials. *Journal of Approximation Theory*, 150(1):69–95, 2008.
- [21] Eli Levin and Doron Lubinsky. Universality limits for exponential weights. *Constructive Approximation*, 29:247–275, 2009.
- [22] Eli Levin and Doron Lubinsky. *Orthogonal polynomials for exponential weights*. Springer Science & Business Media, 2012.
- [23] John Littlewood and Albert Offord. On the number of real roots of a random algebraic equation (iii). *Rec. Math. [Mat. Sbornik]*, 12(3):277–286, 1943.
- [24] John Littlewood and Albert Offord. On the distribution of the zeros and  $\alpha$ -values of a random integral function (i). *Journal of the London Mathematical Society*, 1(3):130–136, 1945.
- [25] John Littlewood and Albert Offord. On the distribution of zeros and  $\alpha$ -values of a random integral function (ii). *Annals of Mathematics*, pages 885–952, 1948.
- [26] Doron Lubinsky and Igor Pritsker. Variance of real zeros of random orthogonal polynomials for varying and exponential weights. *Electronic Journal of Probability*, 27:1–32, 2022.
- [27] Doron Lubinsky, Igor Pritsker, and Xiaoju Xie. Expected number of real zeros for random linear combinations of orthogonal polynomials. *Proceedings of the American Mathematical Society*, 144(4):1631–1642, 2016.
- [28] Doron Lubinsky, Igor Pritsker, and Xiaoju Xie. Expected number of real zeros for random orthogonal polynomials. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 164, 2018.

- [29] Hoi Nguyen and Van Vu. Optimal inverse Littlewood–Offord theorems. *Advances in Mathematics*, 226(6):5298–5319, 2011.
- [30] Hoi. Nguyen and Van. Vu. Small ball probability, inverse theorems, and applications. In *Erdős centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 409–463. János Bolyai Math. Soc., Budapest, 2013.
- [31] Oanh Nguyen and Van Vu. Roots of random functions: A framework for local universality. *American Journal of Mathematics*, 144(1):1–747, 2022.
- [32] Igor Pritsker. Zero distribution of random polynomials. *Journal d’Analyse Mathématique*, 134(2):719–745, 2018.
- [33] Igor Pritsker and Xiaoju Xie. Expected number of real zeros for random Freud orthogonal polynomials. *Journal of Mathematical Analysis and Applications*, 429(2):1258–1270, 2015.
- [34] Walter Rudin. *Real and complex analysis (3rd)*. New York: McGraw-Hill Inc, 1986.
- [35] Edward B Saff and Vilmos Totik. *Logarithmic potentials with external fields*, volume 316. Springer Science & Business Media, 2013.
- [36] Mikhail Sodin. Zeroes of Gaussian analytic functions. In *European congress of mathematics*, pages 445–458, 2005.
- [37] Terence Tao and Van Vu. A sharp inverse Littlewood–Offord theorem. *Random Structures Algorithms*, 37(4):525–539, 2010.
- [38] Terence Tao and Van Vu. Local universality of zeroes of random polynomials. *International Mathematics Research Notices*, page rnu084, 2014.
- [39] J Wilkins Jr. The expected value of the number of real zeros of a random sum of Legendre polynomials. *Proceedings of the American Mathematical Society*, 125(5):1531–1536, 1997.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF VIRGINIA, 141 CABELL DRIVE, CHARLOTTESVILLE, VA 22904, USA

*Email address:* yendo@virginia.edu

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GA 30332, USA

*Email address:* lubinsky@math.gatech.edu

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, 231 W 18TH AVE, COLUMBUS, OH 43210 USA

*Email address:* nguyen.1261@osu.edu

DIVISION OF APPLIED MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02906, USA

*Email address:* oanh\_nguyen1@brown.edu

DEPARTMENT OF MATHEMATICS, OKLAHOMA STATE UNIVERSITY, STILLWATER, OK 74078, USA

*Email address:* igor@math.okstate.edu