Roots of random functions: A framework for local universality

Oanh Nguyen, Van Vu

American Journal of Mathematics, Volume 144, Number 1, February 2022, pp. 1-74 (Article)

Published by Johns Hopkins University Press
DOI: *https://doi.org/10.1353/ajm.2022.0000*

ROOTS OF RANDOM FUNCTIONS: A FRAMEWORK FOR LOCAL UNIVERSALITY

By OANH NGUYEN and VAN VU

Abstract. We investigate the local distribution of roots of random functions of the form $F_n(z) = \sum_{i=1}^n \xi_i \phi_i(z)$, where ξ_i are independent random variables and $\phi_i(z)$ are arbitrary analytic functions. Starting with the fundamental works of Kac and Littlewood-Offord in the 1940s, random functions of this type have been studied extensively in many fields of mathematics.

We develop a robust framework to solve the problem by reducing, via universality theorems, the calculation of the distribution of the roots and the interaction between them to the case where ξ_i are Gaussian. In this special case, one can use the Kac-Rice formula and various other tools to obtain precise answers.

Our framework has a wide range of applications, which include the most popular models of random functions, such as random trigonometric polynomials and all basic classes of random algebraic polynomials (Kac, Weyl, and elliptic). Each of these ensembles has been studied heavily by deep and diverse methods. Our method, for the first time, provides a unified treatment for all of them.

Among the applications, we derive the first local universality result for random trigonometric polynomials with arbitrary coefficients. When restricted to the study of real roots, this result extends several recent results, proved for less general ensembles. For random algebraic polynomials, we strengthen several recent results of Tao and the second author, with significantly simpler proofs. As a corollary, we sharpen a classical result of Erdös and Offord on real roots of Kac polynomials, providing an optimal error estimate. Another application is a refinement of a recent result of Flasche and Kabluchko on the roots of random Taylor series.

1. Introduction. Let n be a positive integer or ∞. Let $\phi_1, \dots, \phi_n$ be deterministic functions and $\xi_1, \dots, \xi_n$ be independent random variables. Consider the random function/series

$$F_n = \sum_{i=1}^n \xi_i \phi_i. \tag{1}$$

A fundamental task is to understand the distribution of and the interaction between the roots (both real and complex) of F_n. For several decades, this task has been carried out in many different areas of mathematics such as analysis, numerical analysis, probability, mathematical physics; see [3, 14, 17, 23, 26, 34, 51, 61], for example.

The most studied subcases are when $\phi_i = c_i x^i$ (in which case F_n is a random algebraic polynomial) and $\phi_i = c_i \cos ix$ (in which case F_n is a random trigonometric polynomial); here and later, the c_i are deterministic coefficients that may

Manuscript received July 7, 2018; revised July 11, 2021.

Research of the first author supported by NSF grant DMS-2125031; research of the second author supported by NSF grant DMS-1307797 and AFORS grant FA9550-12-1-0083.

American Journal of Mathematics 144 (2022), 1–74. © 2022 by Johns Hopkins University Press.

depend on i and n. In fact, these classes split further, according to the values of c_i. For instance, three important classes of random algebraic polynomials are: Kac polynomials ($c_i = 1$), Weyl polynomials ($c_i = \frac{1}{\sqrt{i!}}$) and elliptic polynomials ($c_i = \sqrt{\binom{n}{i}}$). For random trigonometric polynomials, most papers seem to focus on the case $c_i = 1$. A very significant part of the literature on random functions focuses on these special classes.

Even for these classical cases, the problem is already hard; see [1, 2, 10, 19, 29, 30, 44, 53, 54, 58] for a partial list of recent developments. It requires a full book to discuss the results and methods concerning random polynomials, but one feature stands out. The distributions of the roots in different classes are quite different, and the methods to study them are often specialized.

In this paper, we aim to develop a robust framework to solve the general problem. The leading idea is to utilize universality theorems to reduce the problem of calculating the distribution of the roots and the interaction between them to the case where the ξ_i are Gaussian. In the Gaussian case, the answers can be (or, for most ensembles, have already been) computed in a precise form, using the Kac-Rice formula and various other tools which make use of special properties of Gaussian random variables and Gaussian processes; see, for instance [14, 24, 26, 46, 49, 51, 58]. In particular, when the ξ_i are complex Gaussian variables, F_n is called a Gaussian analytic function, and we refer to Sodin's paper [51] for an in-depth survey.

Universality theorems of this type have recently been proved in [11, 58] by the authors, Do and Tao for many classes of random algebraic polynomials of various types, using complex machinery (see also [31, 35, 44, 45, 55] for related works concerning global universality). The method built in these papers is sensitive. It does not apply to random trigonometric polynomials and many other ensembles.

In this paper, we are going to establish a new and general condition which guarantees universality for a wide class of random functions. This class contains all popular random functions. Among others, it covers all classical random algebraic polynomials (such as those considered in [11, 58] and many others). Quite remarkably, it also covers random trigonometric polynomials with general coefficients, whose behavior is totally different. (For readers not familiar with the theory of random functions, let us point out that random trigonometric polynomials typically have $\Theta(n)$ real roots while Kac polynomials have only $\Theta(\log n)$.)

We would like to emphasize the simplicity and robustness of our approach. Proofs of local universality results have been, so far, considerably complex and long. Furthermore, different ensembles require proofs which are different in at least a few key technical aspects. Our proofs, based on new observations, are quite simple and robust. The proof for the general theorem is only a few pages long. Next, and more importantly, we can deduce universality results for completely different ensembles of random functions from this general theorem in an identical way using (essentially) one simply stated lemma. In each ensemble considered, we either obtain completely new results or a short, new proof of the most current result, many times with a quantitative improvement. The length of the paper is due to the number of applications. The reader is invited to read Section 2.4 for a discussion of our method and a comparison with the previous ones.

Let us now briefly discuss the applications. Consider two random functions $F_n = \sum_{i=1}^n \xi_i \phi_i$ and $\tilde{F}_n = \sum_{i=1}^n \tilde{\xi}_i \phi_i$, where ξ_i and $\tilde{\xi}_i$ can have different distributions. We show (under some mild assumptions) that the local statistics of the roots of the two functions are asymptotically the same. In practice, we can set $\tilde{\xi}_i$ to be Gaussian, and thus reduce the study to this case. The local information can be used to derive certain global properties; for instance, the number of roots in a large region (which has been partitioned into many local cells) is simply the sum of the numbers of roots in each cell.

• We study random trigonometric polynomials in Section 3. We derive (to the best of our knowledge) the first local universality of correlation for this class. Our setting is more flexible than most previous works on this topic, as we allow a large degree of freedom in choosing the deterministic coefficients c_i.

While we do not find comparable previous local universality results for random trigonometric polynomials, we can still make some comparisons to previous works by restricting to the popular sub-problem of estimating the density of the real roots. For this problem, our universality result yields new estimates which extend several existing results, some of which are quite recent and have been proved by totally different methods; see Section 3 for details.

• In Section 4, we discuss Kac polynomials. We derive a short proof for a strengthening of a recent result of Tao and the second author [58]. By almost the same argument, one could also recover the main result of Do and the authors [11] which applies for generalized Kac polynomials. As a corollary, we obtain a more precise version of the classical result of Erdös and Offord [15] on the number of real roots.

• In Section 5, we study Weyl series. Our universality result here provides an exact estimate for the expectation of the number of roots in any fixed domain B. Previous to our result, such an estimate was only known for sets of the form rB, where r is a parameter tending to infinity, thanks to a very recent work of Kabluchko and Zaporozhets [30].

• In Section 6, we apply our results to random elliptic polynomials. We give a short proof of a recent result from [58], which generalizes an earlier result of Bleher and Di [6].

• The above applications already cover all traditional classes of random functions in the literature. To illustrate the generality of our result, in Section 7, we present one more application, concerning random series with regularly varying coefficients, a class defined and studied by Flasche and Kabluchko very recently [20].

• While revising this paper, we became aware of a recent work [21] which has some overlaps with ours. We made a brief comparison at the end of Section 7.

• Additionally, after this work had been announced, the framework that we develop here has been applied to the following papers.

– In [8], Mei-Chu Chang, Hoi Nguyen and the authors study the number of intersections between random eigenfunctions of general eigenvalues and a given smooth curve in flat tori.

– In [12], Yen Do and the authors study random orthonormal polynomials.

In most applications, we will work out corollaries concerning the problem of counting real roots. While our results yield much more than just the density function of real roots, we focus on this subproblem since it is, traditionally, one of the most natural and appealing problems in the field. (Technically speaking, the study of zeros of random analytic functions started with papers of Littlewood-Offord and Kac in the 1940s, studying the number of real roots of Kac polynomials.) Our corollaries provide many new contributions to the existing vast literature on this subject. As a matter of fact, our results allow us to study any level set $L_a := \{z \in \mathbb{C} : F_n(z) = a\}$ for any fixed a (the roots form the level set L_0) at no extra cost.

The rest of the paper is organized as follows. In the next section, we first describe our goal, namely, what we mean by universality. We then establish the general condition that guarantees universality, and comment on its strength. We next state the general universality theorems along with a discussion of the main ideas in the proof.

The next 5 sections (Sections 3–7) are devoted to the applications mentioned above. We state universality theorems for various classes of random functions, and derive corollaries concerning the density of both real and complex roots. In Section 8, we prove the general universality theorems stated in Section 2. The rest of the paper is devoted to the verification of the applications in Sections 3–7. We also include a short appendix at the end of the paper, which contains the proofs of a few lemmas (some of which were proved elsewhere), for the sake of completeness.

Acknowledgments. The authors would like to thank Asaf Ferber and Yuval Peres for helpful remarks that led to some simplifications of our proofs. We thank the anonymous referees for their helpful suggestions. Part of this work was done at VIASM (Hanoi), and the authors would like to thank the institute for its support and hospitality.

2. Universality theorems.

In the first subsection, we describe the traditional way to compare local statistics of the roots. Next, we provide the assumptions under which our theorems hold, and comment on their strength. The precise statements come in the final subsection.

The notation $\mathbf{1}_E$ denotes the indicator of an event E; it takes value 1 if E holds and 0 otherwise.

2.1. Comparing local statistics. For simplicity, let us first focus on the complex roots of F_n. These roots form a random point set on the plane.

The first interesting local statistics is the density. In order to understand the density around a point z, we consider the unit disk $B(z,1)$ centered at z. In practice, the radius of the disk is chosen so that the number of roots in it is typically of order $\Theta(1)$. The expected number of roots in the disk can be written as

$$\sum_i \mathbf{E} f(\zeta_i)$$

where $\zeta_1, \zeta_2, \ldots$ are the roots of F_n, and f is the indicator function of $B(z,1)$; in other words, $f(x) = 1$ if $x \in B(z,1)$ and zero otherwise.

If one is interested in the pairwise correlation between the roots near z, then it is natural to look at

$$\sum_{i,j} \mathbf{E} f(\zeta_i, \zeta_j)$$

where $f(x,y)$ is the indicator function of $B(z,1)^2 := B(z,1) \times B(z,1)$; in other words, $f(x,y) = 1$ if both $x, y \in B(z,1)$ and zero otherwise.

In general, the k-wise correlation can be computed from

$$\sum_{i_1,\ldots,i_k} \mathbf{E} f(\zeta_{i_1}, \ldots, \zeta_{i_k})$$

where $f(x_1,\ldots,x_k)$ is the indicator function of $B(z,1)^k$. A good estimate for these quantities tells us how the nearby roots repel or attract each other.

Even more generally, one can study the interaction of roots near different centers by looking at

$$\sum_{i_1,\ldots,i_k} \mathbf{E} f(\zeta_{i_1}, \ldots, \zeta_{i_k})$$

where $f(x_1,\ldots,x_k)$ is the indicator function of $B(z_1,1) \times B(z_2,1) \cdots \times B(z_k,1)$ with $B(z_i,1)$ being the unit disk centered at z_i.

Now, consider another random function

$$\tilde{F}_n = \sum_{i=1}^n \tilde{\xi}_i \phi_i$$

where the $\tilde{\xi}_i$ are independent random variables distributed differently from the ξ_i. We end up with two sets of quantities

$$\sum_{i_1,\ldots,i_k} \mathbf{E} f(\zeta_{i_1}, \ldots, \zeta_{i_k})$$

and

$$\sum_{i_1,\dots,i_k} \mathbf{E} f\big(\tilde\zeta_{i_1},\dots,\tilde\zeta_{i_k}\big)$$

where the $\tilde\zeta_i$ are the roots of $\tilde F_n$.

We would like to show (under certain assumptions) that these two quantities are asymptotically the same, namely

$$\left| \sum_{i_1,\dots,i_k} \mathbf{E} f\big(\zeta_{i_1},\dots,\zeta_{i_k}\big) - \sum_{i_1,\dots,i_k} \mathbf{E} f\big(\tilde\zeta_{i_1},\dots,\tilde\zeta_{i_k}\big) \right| \le \delta_n \tag{2}$$

for some δ_n tending to zero as n goes to infinity.

For technical convenience, we will replace the indicator function f by a smoothed approximation. This makes no difference in applications. On the other hand, our results hold for any smoothed test function f, which may have nothing to do with the indicator function.

If one cares about the real roots, one replaces the disk $B(z,1)$ by the interval of length 1 centered at a real number z. In general, instead of the product $B(z_1,1)\times B(z_2,1)\cdots\times B(z_k,1)$, one can consider a mixed product of disks and intervals. This enables one to understand the interaction between nearby roots of both types (complex and real).

One, of course, could have made the previous discussion using the notion of correlation functions. However, we find the current format direct and intuitive. We refer to [26] or [58] for more detailed discussions concerning local statistics using correlation functions.

2.2. Assumptions. Before stating the result, let us discuss the assumptions. There are two types of assumptions. The first is for the random variables ξ_i and $\tilde\xi_i$. The second concerns the deterministic functions ϕ_i.

For the random variables, our assumption is close to minimal. In the case that both ξ_i and $\tilde\xi_i$ are real, our simplest assumption is

Condition C0. The random variables $\xi_1,\dots,\xi_n,\tilde\xi_1,\dots,\tilde\xi_n$ are independent real random variables with the same mean $\mathbf{E}\xi_i = \mathbf{E}\tilde\xi_i$ for each i, variance one, and (uniformly) bounded $(2+\varepsilon)$ central moments, for some constant $0<\varepsilon<1$.

In fact, we can relax the assumption of matching means and variances, allowing a finite number of exceptions. If the ξ_i and $\tilde\xi_i$ are complex, the matching mean and variance need to be adjusted to address both real and imaginary parts.

Condition C1. Two sequences of random variables $(\xi_1,\dots,\xi_n)$ and $(\tilde\xi_1,\dots,\tilde\xi_n)$ are said to satisfy this condition if the following hold, for some constants $N_0,\tau>0$ and $0<\varepsilon<1$.

(i) *Uniformly bounded* $(2+\varepsilon)$ *central moments:* The random variables ξ_i (and similarly $\tilde{\xi}_i$), $1 \le i \le n$, are independent (real or complex, not necessarily identically distributed) random variables with unit variance (namely, $\mathbf{E}|\xi_i - \mathbf{E}\xi_i|^2 = 1$), and bounded $(2+\varepsilon)$ central moments, namely $\mathbf{E}|\xi_i - \mathbf{E}\xi_i|^{2+\varepsilon} \le \tau$.

(ii) *Matching moments to second order with finitely many exceptions:* For any $i \ge N_0$, for all $a, b \in \{0,1,2\}$ with $a+b \le 2$,

$$\mathbf{E}\,\mathrm{Re}\left(\xi_i\right)^a \mathrm{Im}\left(\xi_i\right)^b = \mathbf{E}\,\mathrm{Re}\left(\tilde{\xi}_i\right)^a \mathrm{Im}\left(\tilde{\xi}_i\right)^b,$$

and for $0 \le i < N_0$, $\left|\mathbf{E}\xi_i - \mathbf{E}\tilde{\xi}_i\right| \le \tau$.

It is trivial that Condition C1 contains Condition C0 as a special case. We find it rewarding to go with the more general, but slightly technical, assumption (ii), which allows non-matching means, as it leads to an interesting phenomenon that changing a finite number of terms in $F_n(z)$ does not influence the asymptotic distribution of the roots. Among other benefits, this allows us to generalize all results to level sets $\{z \in \mathbb{C} : F_n(z) = a\}$ for any fixed a; see Remark 3.7 for more details.

We now turn to the assumption on the deterministic functions ϕ_i. The statement of our theorems will involve two parameters, an error term $0 < \delta_n < 1$ (see (2)) and a region $D_n \subset \mathbb{C}$, from which the base points $z_1, \dots, z_k$ are chosen. As their subscripts indicate, both δ_n and D_n can depend on n. In most of our applications, δ_n tends to zero with n but it is not required. When $n = \infty$ for example, δ_∞ can be any parameter in $(0,1)$. The assumptions below are tailored to these two parameters, δ_n and D_n.

For two sets $\mathcal{A}, \mathcal{B} \subset \mathbb{C}$, define $\mathcal{A} + \mathcal{B} := \{a + b : a \in \mathcal{A}, b \in \mathcal{B}\}$. Let $k, C_1, \alpha_1, A, c_1, C$ be positive constants. We say that F_n satisfies Condition C2 with parameters $(k, C_1, \alpha_1, A, c_1, C)$ if the following holds.

Condition C2. (1) For any $z \in D_n$, F_n is analytic on the disk $B(z,2)$ with probability 1 and

$$\mathbf{E} N^{k+2} \mathbf{1}_{N \ge \delta_n^{-C_1}} \le C,$$

where N is the number of zeros of F_n in the disk $B(z,1)$. We note that throughout this paper, if F_n is identically 0, we adopt the (admittedly artificial) convention that F_n has no roots in $\mathbb{C}$.

(2) *Anti-concentration:* For every $z \in D_n$, with probability at least $1 - C\delta_n^A$, there exists $z' \in B(z, 1/100)$ for which $|F_n(z')| \ge \exp(-\delta_n^{-c_1})$.

(3) *Boundedness:* For any $z \in D_n$, with probability at least $1 - C\delta_n^A$, $|F_n(w)| \le \exp(\delta_n^{-c_1})$ for all $w \in B(z,2)$.

(4) *Delocalization:* For every $z \in D_n + B(0,1)$, it holds that $\sum_{j=1}^n |\phi_j(z)|^2 \neq 0$ and for every $i = 1, \dots, n$,

$$\frac{|\phi_i(z)|}{\sqrt{\sum_{j=1}^n |\phi_j(z)|^2}} \leq C\delta_n^{\alpha_1}.$$

(5) *Derivative growth:* For any real number $x \in D_n + B(0,1)$,

$$\sum_{j=1}^n |\phi_j'(x)|^2 \leq C\delta_n^{-c_1} \sum_{j=1}^n |\phi_j(x)|^2,$$

$$\sum_{j=1}^n \sup_{z \in B(x,1)} |\phi_j''(z)|^2 \leq C\delta_n^{-c_1} \sum_{j=1}^n |\phi_j(x)|^2,$$

and

$$\sum_{j=1}^n |\mathbf{E}\xi_j| \sup_{z \in B(x,1)} |\phi_j''(z)| \leq C\delta_n^{-c_1} \sqrt{\sum_{j=1}^n |\phi_j(x)|^2}.$$

Remark 2.1. While Condition C2 still involves the random variables ξ_i, in the verification of these conditions, we only need to use basic information about the mean of these variables. On the other hand, the type of arguments one needs to use in the verification depends strongly on the functions ϕ_i.

Remark 2.2. The last Condition C2(5) is important only in the study of real roots; in particular, it is used to prove the repulsion of the real roots (Lemma 8.5). It can be ignored in the study of complex roots.

Let us now comment on the verification of Condition C2 in practice.

Remark 2.3. Typically, we assume δ_n tends to zero with n. We transform the functions so that the expectation of N is of order 1 where N is the number of roots of F_n in a disk $B(z,1)$, $z \in D_n$. With this in mind, the first condition is a large deviation estimate on N and can be proved using standard large deviation tools combined with classical complex analytic estimates such as Jensen's inequality. The third condition (boundedness) is also a large deviation statement and can be dealt with using standard tools, since for any fixed w, $F_n(w)$ is a sum of independent random variables.

The two Conditions C2(4) and C2(5) are deterministic properties of the functions ϕ_i and hold for many natural classes of functions. The forth condition (delocalization) simply says that in the vector $(\phi_i(z))_1^n$, no coordinate dominates. The fifth condition asserts that the first and second derivatives of ϕ_i do not exceed the value of the function itself by a large multiplicative factor, in an average sense.

Checking these conditions is usually a routine task. Furthermore, the proof allows us to easily modify these conditions, if necessary.

The second (anti-concentration) condition is the one that may require some work. However, this condition is trivial if (some of) the random variables ξ_i have continuous distributions with bounded density. For instance, if $\phi_1 = 1$ (constant function) and ξ_1 has a continuous distribution with bounded density, then the required anti-concentration property holds trivially by conditioning on the rest of the random variables (which can have arbitrary distributions). There is a sizable literature focusing on continuous ensembles, and our results allow us to recover, in a straightforward manner, a number of existing results, whose original proofs were quite technical; see Sections 4 and 6 for examples.

2.3. Results.

Given the assumptions discussed in the previous section, we are now ready to state our universality theorems.

Definition 2.4. For any function $G : \mathbb{R}^k \to \mathbb{R}$ and any natural number a, we define $\|\nabla^a G\|_\infty$ to be the supremum over $x \in \mathbb{R}^k$ of the absolute value of all partial derivatives of total order a of G at x. For a function $G : \mathbb{R}^k \times \mathbb{C}^l \to \mathbb{C}$, we define $\|\nabla^a G\|_\infty$ to be the maximum of $\|\nabla^a G_1\|_\infty$ and $\|\nabla^a G_2\|_\infty$, where $G_1, G_2 : \mathbb{R}^{k+2l} \to \mathbb{R}$ are the real and imaginary parts of G:

$$G_1(x_1,\dots,x_k,u_1,\dots,u_l,v_1,\dots,v_l) = \mathrm{Re}\,\big(G(x_1,\dots,x_k,u_1+iv_1,\dots,u_l+iv_l)\big),$$
$$G_2(x_1,\dots,x_k,u_1,\dots,u_l,v_1,\dots,v_l) = \mathrm{Im}\,\big(G(x_1,\dots,x_k,u_1+iv_1,\dots,u_l+iv_l)\big).$$

THEOREM 2.5. (General complex universality) *Assume that the coefficients ξ_i and $\tilde{\xi}_i$ satisfy Condition C1 for some constants N_0, τ, ε. Let α_1, C_1 be positive constants and k be a positive integer. Set $A := 2kC_1 + \frac{\alpha_1\varepsilon}{60}$ and $c_1 := \frac{\alpha_1\varepsilon}{10^5k^2}$. Assume that there exists a constant $C > 0$ such that the random functions F_n and $\tilde{F}_n$ satisfy Conditions C2(1)–C2(4) with parameters $(k, C_1, \alpha_1, A, c_1, C)$. Then there exist positive constants C', c depending only on the constants in Conditions C1 and C2 (but not on δ_n, D_n and n) such that the following holds.*

For any complex numbers $z_1,\dots,z_k$ in D_n and any function $G : \mathbb{C}^k \to \mathbb{C}$ supported on $\prod_{i=1}^k B(z_i, 1/100)$ with continuous derivatives up to order $2k+4$ and $\|\nabla^a G\|_\infty \le 1$ for all $0 \le a \le 2k+4$, we have

$$\left|\mathbf{E}\sum G(\zeta_{i_1},\dots,\zeta_{i_k}) - \mathbf{E}\sum G(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k})\right| \le C'\delta_n^c, \tag{3}$$

where the first sum runs over all k-tuples $(\zeta_{i_1},\dots,\zeta_{i_k})$ of the roots $\zeta_1,\zeta_2,\dots$ of F_n, and the second sum runs over all k-tuples $(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k})$ of the roots $\tilde{\zeta}_1,\tilde{\zeta}_2,\dots$ of $\tilde{F}_n$.

As an example for the summation in (3), if $k = 2$ and F_n only has two roots ζ_1 and ζ_2, then the first sum is $G(\zeta_1,\zeta_1) + G(\zeta_1,\zeta_2) + G(\zeta_2,\zeta_1) + G(\zeta_2,\zeta_2)$.

THEOREM 2.6. (General real universality) *Assume that* $\phi_i(\mathbb{R}) \subset \mathbb{R}$ *and* ξ_i *and* $\tilde{\xi}_i$ *are real random variables that satisfy Condition C1 for some constants* N_0, τ, ε. *Let* α_1, C_1 *be positive constants and* k, l *be nonnegative integers with* $k+l \geq 1$. *Set* $A = 2(k+l+2)(C_1+2) + \frac{\alpha_1 \varepsilon}{60}$ *and* $c_1 = \frac{\alpha_1 \varepsilon}{10^9 (k+l)^4}$. *Assume that there exists a constant* $C > 0$ *such that the random functions* F_n *and* $\tilde{F}_n$ *satisfy Condition C2 with parameters* $(k+l, C_1, \alpha_1, A, c_1, C)$. *Then there exist positive constants* C', c *depending only on* k, l *and the constants in Conditions C1 and C2 (but not on* δ_n, D_n *and* n*) such that the following holds.*

For any real numbers $x_1, \dots, x_k$, *complex numbers* $z_1, \dots, z_l$, *all of which are in* D_n, *and any function* $G : \mathbb{R}^k \times \mathbb{C}^l \to \mathbb{C}$ *supported on* $\prod_{i=1}^k [x_i - 1/100, x_i + 1/100] \times \prod_{j=1}^l B(z_j, 1/100)$ *with continuous derivatives up to order* $2(k+l)+4$ *and* $\|\nabla^a G\|_\infty \leq 1$ *for all* $0 \leq a \leq 2(k+l)+4$, *we have*

$$\left| \mathbf{E} \sum G(\zeta_{i_1}, \dots, \zeta_{i_k}, \zeta_{j_1}, \dots, \zeta_{j_l}) - \mathbf{E} \sum G(\tilde{\zeta}_{i_1}, \dots, \tilde{\zeta}_{i_k}, \tilde{\zeta}_{j_1}, \dots, \tilde{\zeta}_{j_l}) \right| \leq C' \delta_n^c,$$

where the first sum runs over all $(k+l)$*-tuples* $(\zeta_{i_1}, \dots, \zeta_{i_k}, \zeta_{j_1}, \dots, \zeta_{j_l}) \in \mathbb{R}^k \times \mathbb{C}_+^l$ *of the roots* $\zeta_1, \zeta_2, \dots$ *of* F_n, *and the second sum runs over all* $(k+l)$*-tuples* $(\tilde{\zeta}_{i_1}, \dots, \tilde{\zeta}_{i_k}, \tilde{\zeta}_{j_1}, \dots, \tilde{\zeta}_{j_l}) \in \mathbb{R}^k \times \mathbb{C}_+^l$ *of the roots* $\tilde{\zeta}_1, \tilde{\zeta}_2, \dots$ *of* $\tilde{F}_n$.

Remark 2.7. The specific values of A and c_1 in both theorems are chosen for the sake of explicitness. The theorems hold for any bigger A and any smaller c_1. The constant c in both theorems can be chosen to be c_1, namely $\frac{\alpha_1 \varepsilon}{10^5 k^2}$ and $\frac{\alpha_1 \varepsilon}{10^9 (k+l)^4}$, respectively. We make no attempt to optimize these constants.

2.4. Main ideas and technical novelties.

2.4.1. Main ideas. Let us consider the simplest setting where $k = 1$, $l = 0$ and we need to show

$$\sum_{i=1}^n \mathbf{E} G(\zeta_i) = \sum_{i=1}^n \mathbf{E} G(\tilde{\zeta}_i) + O(\delta_n^c),$$

where the ζ_i (and the $\tilde{\zeta}_i$) are the roots of F_n (and $\tilde{F}_n$, respectively) and G is a (smooth) test function supported on a disk $B(z_0, 1/100)$.

Our starting point is the Green's formula, which asserts that

$$G(0) = \frac{1}{2\pi} \int_{\mathbb{C}} \log |z| \Delta G(z) dz.$$

By change of variables, this implies that for all i,

$$G(\zeta_i) = \frac{1}{2\pi} \int_{\mathbb{C}} \log |z - \zeta_i| \Delta G(z) dz,$$

which, in turn, yields

$$\begin{aligned}\sum_i \mathbf{E}G(\zeta_i) &= \frac{1}{2\pi}\mathbf{E}\int_{\mathbb{C}} \log\left|\prod_{i=1}^{n}(z-\zeta_i)\right|\Delta G(z)dz\\ &= \frac{1}{2\pi}\mathbf{E}\int_{B(z_0,1/100)} \log\left|F_n(z)\right|\Delta G(z)dz.\end{aligned}$$

An obvious, and major, technical difficulty here is that the logarithmic function has a singularity at 0. This, naturally, leads to the anti-concentration issue that we discussed earlier, namely we need to bound the probability that $|F_n(z)|$ is close to zero. Condition C2(2) has been introduced to address this issue.

Let us assume, for a moment, that the singularity problem has been handled properly (we will discuss the anti-concentration property shortly). Then, by using Conditions C2(1)–C2(3), we can show that the function F_n is nice enough that we can replace $\log|F_n|$ by $K(F_n)$ where K is a bounded smooth function. The key argument of this part is to bound the error term, which turns out to be relatively simple.

The task is now reduced to showing that

$$\mathbf{E}\int_{B(z_0,1/100)} K\big(F_n(z)\big)\Delta G(z)dz - \mathbf{E}\int_{B(z_0,1/100)} K\big(\tilde{F}(z)\big)\triangle G(z)dz = O\big(\delta^c\big).$$

Because of the boundedness of G, for each $z \in B(z_0, 1/100)$, it suffices to show that

$$\mathbf{E}K\big(F_n(z)\big) - \mathbf{E}K\big(\tilde{F}(z)\big) = O\big(\delta^c\big).$$

Since for each fixed z, $F_n(z)$ is a sum of independent random variables, the desired bound can be viewed, in some sense, as a quantitative version of the Central Limit Theorem. We will actually prove it by the Lindeberg swapping method, which, by now, is a standard tool for proving local universality.

Generalizing the whole scheme to the general case of k and l requires several additional technical steps, but the spirit of the method remains the same.

2.4.2. Comparison with earlier papers [11, 58]**.** Our method differs from that of [58] at essential steps. The first key idea in [58] is to handle the integral

$$\frac{1}{2\pi}\mathbf{E}\int_{B(z_0,1/100)} \log\left|F_n(z)\right|\Delta G(z)dz$$

by a random Riemann sum. One tries to approximate this integration by $\frac{c}{m}(f(z_1) + \cdots + f(z_m))$, where z_i are iid random points sampled from the disk, m is a properly chosen parameter which tends to infinity with n, c is a normalizing constant, and $f := \log|F_n|\Delta G$.

With that approach, one faces two major technical tasks. The first (and harder one) is to control the error term in the approximation. This leads to the problem of estimating the variance in the sampling process. The other task is to prove a comparison estimate for the random vector $(f(z_1), \dots, f(z_m))$, where we now view the points $z_1, \dots, z_m$ as fixed, with the randomness coming from $F_n(z)$. This, again, can be done using a Lindeberg type argument (applying to high dimensional setting).

Our new proof avoids this sampling argument completely, making the argument much shorter and more direct. For instance, the proof of Theorem 2.5, barring some lemmas in the appendix, is now only 3 pages.

Let us now discuss the critical anti-concentration property. In practice, it has been a major issue to prove that a random function satisfies the anti-concentration phenomenon in some way. (As pointed out earlier, this is needed in order to address the singularity problem concerning the logarithmic function.)

In earlier papers [58, 11], every class of random (algebraic) polynomials required a different proof. In [58], for Weyl and elliptic polynomials, the authors used Littlewood-Offord arguments for lacunary sequences. In the same paper, the proof for Kac polynomials required a much more sophisticated argument, based on the Inverse Littlewood-Offord theory (see Nguyen-Vu [42]) and a weak version of the quantitative (Gromov) rigidity theorem (see Shalom-Tao [50]). However, this proof does not hold for the derivatives of Kac polynomials and random polynomials with slowly growing coefficients. In order to handle these classes, in [11], the authors needed to use a beautiful result on log-integrability by Nazarov-Nishry-Sodin [39], a very recent development. However, none of these tools works for random trigonometric polynomials, whose roots behave quite differently.

An important new point in our proof is that we require a much weaker anti-concentration property than in previous papers. We only require that $F_n(z)$, as a random variable, satisfies the anti-concentration for only one point z in the whole neighborhood, while in [58] one requires anti-concentration to hold for most points in the same neighborhood. (Notice that since we are taking an integration with respect to z, this earlier requirement from [58] looks natural.) The key to this observation is our Lemma 8.2, which asserts that under favorable conditions, a lower bound on $|F_n(w)|$ guarantees a weaker, but still useful, lower bound for $|F_n(z)|$ for any z in a neighborhood of w.

Building upon this new observation, we have developed a novel method (based on old results of Turán and Halász) to verify the anti-concentration property in a simple and robust manner. This effort leads to Lemma 9.2, which we can use, in a rather straightforward way, to prove the desired anti-concentration property for all ensembles of random functions discussed in this paper (including all the algebraic polynomials discussed above, random trigonometric polynomials with general coefficients, and a very recent ensemble studied by Flasche-Kabluchko [20]).

3. Application: Universality for random trigonometric polynomials. In this section, we apply our theorems to study *random trigonometric polynomials* of the following form

$$P_n(x) = \sum_{j=0}^{n} c_j \xi_j \cos(jx) + \sum_{j=1}^{n} d_j \eta_j \sin(jx)$$

where c_j and d_j are deterministic coefficients, and $\xi_0, \xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_n$ are independent random variables with unit variance. All of the c_j, d_j, ξ_j and η_j may depend on n.

Most of the existing literature deals with the special case $c_i = d_i = 1$ or $c_i = 1$, $d_i = 0$ for every i. The generality of our study enables us to consider more general coefficients. All we need to assume about the coefficients c_i, d_i is the following:

Condition C3. There exist positive constants τ_1, c and an interval $\mathcal{I}_0 \subset \{1, \dots, n\}$ of size at least cn such that

$$|c_i| \ge \tau_1 \max_{0 \le j \le n} \{|c_j|, |d_j|\} \quad \text{for all } i \in \mathcal{I}_0. \tag{4}$$

With regard to the random variables, we assume that they have mean 0, except for finitely many of them whose mean can be as large as $n^{1/2+o(1)}$. Specifically, we assume

Condition C4. There is a constant $N_0 \ge 0$ such that for $i \ge N_0$, $\mathbf{E}\xi_i = \mathbf{E}\eta_i = 0$ and for $0 \le i < N_0$, $|\mathbf{E}\xi_i| \le n^{\tau_0}$, and $|\mathbf{E}\eta_i| \le n^{\tau_0}$, where $\tau_0 := 1/2 + 10^{-11}\varepsilon$.

The ε in this condition is the ε in Condition C1. The constant τ_0 is not optimal but we make no attempt to improve it. We use the same notation N_0 in both Condition C4 and Condition C1, as we can always replace two different N_0 by their maximum. The assumption that $\mathcal{I}_0$ is an interval is only used in the following simple lemma.

LEMMA 3.1. *Let $\mathcal{I}_0$ be an interval in $\{1, \dots, n\}$ of length βn, for some constant $\beta > 0$. Then there is a constant $\beta' > 0$ such that for any real number a, the set $\mathcal{I}_0$ contains a subset J_a of size at least $\beta' n$, where* $\min_{k \in \mathbb{Z}}\{|2aj - (2k+1)\pi|\} \ge \beta'$ *for all $j \in J_a$.*

Let

$$\tilde{P}_n(x) = \sum_{j=0}^{n} c_j \tilde{\xi}_j \cos(jx) + \sum_{j=1}^{n} d_j \tilde{\eta}_j \sin(jx)$$

where $\tilde{\xi}_0, \tilde{\xi}_1, \dots, \tilde{\xi}_n$ and $\tilde{\eta}_1, \dots, \tilde{\eta}_n$ are some other independent random variables.

THEOREM 3.2. (Universality for trigonometric polynomials) *Let k, l be non-negative integers. Assume that the real coefficients c_i and d_i satisfy Condition*

C3 and the two sequences of real random variables $(\xi_0,\dots,\xi_n,\eta_1,\dots,\eta_n)$ *and* $(\tilde{\xi}_0,\dots,\tilde{\xi}_n,\tilde{\eta}_1,\dots,\tilde{\eta}_n)$ *satisfy Conditions C1 and C4. Then for any positive constant C, there exist positive constants* C',c *depending only on* C,k,l *and the constants in Conditions C1, C3, and C4 such that the following holds.*

For any real numbers $x_1,\dots,x_k$*, and complex numbers* $z_1,\dots,z_l$ *such that* $|\mathrm{Im}(z_j)|\le C/n$ *for all* $1\le j\le l$*, and for any function* $G:\mathbb{R}^k\times\mathbb{C}^l\to\mathbb{C}$ *supported on* $\prod_{i=1}^k[x_i-1/n,x_i+1/n]\times\prod_{j=1}^l B(z_j,1/n)$ *with continuous derivatives up to order* $2(k+l)+4$ *and* $\|\nabla^a G\|_\infty\le n^a$ *for all* $0\le a\le 2(k+l)+4$*, we have*

$$\left|\mathbf{E}\sum G(\zeta_{i_1},\dots,\zeta_{i_k},\zeta_{j_1},\dots,\zeta_{j_l})-\mathbf{E}\sum G(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k},\tilde{\zeta}_{j_1},\dots,\tilde{\zeta}_{j_l})\right|\le C'n^{-c},$$

where the first sum runs over all $(k+l)$*-tuples* $(\zeta_{i_1},\dots,\zeta_{i_k},\zeta_{j_1},\dots,\zeta_{j_l})\in\mathbb{R}^k\times\mathbb{C}_+^l$ *of the roots* $\zeta_1,\zeta_2,\dots$ *of* P_n*, and the second sum runs over all* $(k+l)$*-tuples* $(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k},\tilde{\zeta}_{j_1},\dots,\tilde{\zeta}_{j_l})\in\mathbb{R}^k\times\mathbb{C}_+^l$ *of the roots* $\tilde{\zeta}_1,\tilde{\zeta}_2,\dots$ *of* $\tilde{P}_n$*.*

To the best of our knowledge, the above theorems seem to be the first universality results concerning local statistics of the roots of random trigonometric polynomials. To make a comparison to existing literature, let us focus on the distribution of real roots, which is the case $k=1,l=0$ in Theorem 3.2.

The number of real roots has been a main focus of the study of random trigonometric polynomials. The Gaussian setting has been investigated by a number of researchers, including Dunnage [13], Sanbandham [48], Das [9], Wilkins [60], Edelman and Kostlan [14] and many others. One can compute an exact answer for the expectation using either Kac-Rice formula or Edelman-Kostlan formula [14].

For the non-Gaussian case, little has been known until very recently. Angst and Poly [1], in a recent preprint, proved the asymptotics of the mean number of roots of P_n in a fixed interval $[a,b]$ under the assumptions of finite fifth moment and a Cramer-type condition. Their approach introduced a novel way to work with the Kac-Rice formula which had been considered to be difficult in discrete settings. Using an approach originated by Erdös-Offord [15] and later developed by Ibragimov-Maslova [27] [28], Flasche [19] extended the result in [1] with assumptions on the first two moments only. Let $N_{P_n}(a,b)$ denote the number of real roots of P_n in an interval $[a,b]$.

THEOREM 3.3. (Flasche [19]) *Let* $u\in\mathbb{R}$ *and* $0\le a<b\le 2\pi$ *be fixed numbers. Let* $P_n(x)=u\sqrt{n}+\sum_{j=0}^n\xi_j\cos(jx)+\sum_{j=1}^n\eta_j\sin(jx)$ *where* ξ_j *and* η_j*,* $j\in\mathbb{N}$*, are iid random variables with mean 0 and variance 1. Then*

$$\lim_{n\to\infty}\frac{\mathbf{E}N_{P_n}(a,b)}{n}=\frac{b-a}{\pi\sqrt{3}}\exp\left(-\frac{u^2}{2}\right).$$

Notice that in this theorem, the interval $[a,b]$ contains a linear number of roots. For smaller intervals, a few years ago, Azaïs and coauthors [2] showed that if ξ_i and η_i are iid with a smooth density function, then in an interval of size $\Theta(1/n)$, the number of real zeros converges in distribution to that of a suitable Gaussian process (and is thus universal). In an even more recent paper [29], Iksanov-Kabluchko-Marynych removed the assumption of smooth density, using a different method.

THEOREM 3.4. (Iksanov-Kabluchko-Marynych [29]) *Let*

$$P_n(x) = \sum_{j=0}^{n} \xi_j \cos(jx) + \sum_{j=1}^{n} \eta_j \sin(jx)$$

where (ξ_j, η_j), $j \in \mathbb{N}$, *are iid real random vectors with mean* 0 *and unit covariance matrix. Let* (s_n) *be any sequence of real numbers and* $[a,b] \subset \mathbb{R}$ *a fixed interval. Then*

$$N_{P_n}\left(s_n + \frac{a}{n}, s_n + \frac{b}{n}\right) \xrightarrow[n\to\infty]{d} N_Z(a,b)$$

where $(Z(t))_{t\in R}$ *is the stationary Gaussian process with mean 0 and covariance matrix*

$$\operatorname{Cov}\big(Z(t), Z(s)\big) = \begin{cases} \dfrac{\sin(t-s)}{t-s} & \text{if } t \neq s \\ 1 & \text{if } t = s. \end{cases}$$

In all of these previous works, the coefficients c_i, d_i are: $c_i = d_i = 1$ or $c_i = 1$, $d_i = 0$. Our setting is more general, as we only require a linear fraction of the c_i to be sufficiently large and allow the rest of the (smaller) coefficients to be arbitrary.

Our result implies the following corollary concerning the number of real roots.

THEOREM 3.5. *Under the assumptions of Theorem 3.2, there exist positive constants* C *and* c *such that for any* n *and for any numbers* $a_n < b_n$, *we have*

$$\frac{\left|\mathbf{E}N_{P_n}(a_n,b_n) - \mathbf{E}N_{\tilde{P}_n}(a_n,b_n)\right|}{(b_n-a_n)n} \le Cn^{-c}\left(1 + \frac{1}{(b_n-a_n)n}\right).$$

By using the Kac-Rice formula (Proposition 10.1) for the Gaussian case, we obtain the following precise estimate.

COROLLARY 3.6. *Let* C, ε *and* τ_1 *be positive constants. Let* $-C \le u_n \le C$ *be a deterministic number. Let*

$$P_n(x) = u_n \sqrt{\sum_{i=0}^{n} c_i^2} + \sum_{j=0}^{n} c_j \xi_j \cos(jx) + \sum_{j=1}^{n} c_j \eta_j \sin(jx)$$

where ξ_j *and* η_j, $j \le n$, *are independent (not necessarily identically distributed) real random variables with mean 0, variance 1 and* $(2+\varepsilon)$*-moments bounded by*

C, and the real coefficients c_j satisfy condition C3. Then for any numbers $a_n < b_n$, we have

$$\mathbf{E}N_{P_n}(a_n,b_n) = \frac{b_n-a_n}{\pi}\sqrt{\frac{\sum_{j=0}^n c_j^2 j^2}{\sum_{j=0}^n c_j^2}}\exp\left(-\frac{u_n^2}{2}\right)+O\big(n^{-c}\big)\big((b_n-a_n)n+1\big)$$

where the positive constant c and the implicit constant depend only on C,ε and τ_1.

This corollary extends both Theorems 3.3 and 3.4 in the sense that it holds for general coefficients c_i, d_i and intervals of all scales. It does not seem that the methods used in these papers can cover the same range. On the other hand, our random coefficients are required to have bounded $(2+\varepsilon)$-moments. It is an interesting open problem to see to what extent this assumption is necessary.

Remark 3.7. In the proof, we will show that Corollary 3.6 holds for a more general case in which

$$(5)\qquad \begin{aligned} P_n(x) = {} & \sqrt{\sum_{i=0}^n c_i^2}\left(u_n+\sum_{j=0}^{N_0} u_j n^{-\alpha}\cos(jx)+\sum_{j=1}^{N_0} v_j n^{-\alpha}\sin(jx)\right) \\ & +\sum_{j=0}^n c_j\xi_j\cos(jx)+\sum_{j=1}^n c_j\eta_j\sin(jx) \end{aligned}$$

where $N_0, \alpha > 0$ are any constants and $-C \le u_j, v_j \le C$ are deterministic numbers that can depend on n. This means that the result is applicable to not only the number of zeros of P_n but also the number of intersections between P_n and a deterministic trigonometric polynomial

$$Q(x) := \sqrt{\sum_{i=0}^n c_i^2}\left(u_n'+\sum_{j=0}^{N_0} u_j n^{-\alpha}\cos(jx)+\sum_{j=1}^{N_0} v_j n^{-\alpha}\sin(jx)\right)$$

where u_n', u_j and v_j are bounded deterministic numbers. To see this, one only needs to apply the result to the random polynomial $P_n - Q$.

Now let us go back to the special case with $c_i = d_i = 1$

$$P_n(x) = \sum_{i=0}^n \xi_i\cos(ix)+\sum_{i=1}^n \eta_i\sin(ix).$$

By applying Corollary 3.6 directly to the derivatives of P_n, we get the following result.

COROLLARY 3.8. *Let k be a nonnegative integer and C be a positive constant. Assume that the random variables ξ_i and η_i, $i \le n$, are independent (not necessarily* *identically distributed) real random variables with mean 0, variance 1 and* $(2+\varepsilon)$*-moments bounded by* C*. For any numbers* $a_n < b_n$*, the expected number of real zeros of the kth derivative of* P_n *in an interval* $[a_n, b_n]$ *is*

$$\mathbf{E}N_{P_n^{(k)}}(a_n, b_n) = \sqrt{\frac{2k+1}{2k+3}}\frac{(b_n - a_n)n}{\pi} + O\big(n^{-c}\big)\big((b_n - a_n)n + 1\big)$$

where the positive constant c *and the implicit constant depend only on* k, C *and* ε*.*

The key to our proof is the new technique to verify anti-concentration, which we discussed at the end of the Introduction (see also Remark 2.3) and at the end of the previous section. For details, see Section 9.

4. Application: Universality for Kac polynomials. In this section, we apply our result to Kac polynomials,

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

where $\xi_0, \xi_1, \ldots, \xi_n$ are iid copies of a real random variable ξ with mean zero and unit variance. This is perhaps the most studied model of random polynomials. Indeed, the starting point of the theory of random functions was a series of papers in the early 1900s examining the number of real roots of the Kac polynomials.

The first rigorous work on random polynomials was due to Bloch and Polya in 1932 [7], who considered the Kac polynomial with ξ being Rademacher, namely $\mathbf{P}(\xi = 1) = \mathbf{P}(\xi = -1) = 1/2$. In what follows, we denote by $N_{n,\xi}$ the number of real roots of $P_n(x)$. Next came the ground-breaking series of papers by Littlewood and Offord [37, 38, 36] in the early 1940s, which, to the surprise of many mathematicians at the time, showed that $N_{n,\xi}$ is typically poly-logarithmic in n.

THEOREM 4.1. (Littlewood-Offord) *For* ξ *being Rademacher, Gaussian, or uniform on* $[-1, 1]$*,*

$$\frac{\log n}{\log\log n} \leq N_{n,\xi} \leq \log^2 n$$

with probability $1 - o(1)$*.*

During more or less the same time, Kac [32] discovered his famous formula for the density function $\rho(t)$ of $N_{n,\xi}$

$$\rho(t) = \int_{-\infty}^{\infty} |y| p(t, 0, y)\, dy,$$

where $p(t, x, y)$ is the joint probability density of $P_n(t) = x$ and the derivative $P_n'(t) = y$.

Consequently,

$$\mathbf{E}N_{n,\xi} = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} |y| p(t,0,y) dy. \tag{6}$$

In the Gaussian case (ξ is Gaussian), one can compute the joint distribution of $P_n(t)$ and $P_n'(t)$ rather easily. Kac showed in [32] that

$$\mathbf{E}N_{n,\text{Gauss}} = \frac{1}{\pi}\int_{-\infty}^{\infty}\sqrt{\frac{1}{(t^2-1)^2}+\frac{(n+1)^2t^{2n}}{(t^{2n+2}-1)^2}}dt = \left(\frac{2}{\pi}+o(1)\right)\log n.$$

In his original paper [32], Kac thought that his formula would lead to the same estimate for $\mathbf{E}N_{n,\xi}$ for all other random variables ξ. It has turned out not to be the case, as the right-hand side of (6) is often hard to compute, especially when ξ is discrete (Rademacher for instance). Technically, the computation of the joint distribution of $P_n(t)$ and $P_n'(t)$ is easy in the Gaussian case, thanks to special properties of the Gaussian distribution, but can pose a great challenge in general. Kac admitted this in a later paper [33] in which he managed to push his method to treat the case ξ being uniform in $[-1,1]$, using analytic tools. A further extension was made by Stevens [56], who evaluated Kac's formula for a large class of ξ having continuous and smooth distributions with certain regularity properties (see [56, p. 457] for details). Since the distributions are smooth, the two later results follow rather easily from our universality results; see the discussion at the end of the last section and Remark 2.3; we leave the routine verification as an exercise for the interested reader.

The computation of $\mathbf{E}N_{n,\xi}$ for discrete random variables ξ required a considerable effort. It took more than 10 years until Erdös and Offord [15] found a completely new approach to handle the Rademacher case, proving the following.

THEOREM 4.2. [15] *Let ξ_i be iid Rademacher random variables. Then*

$$N_{n,\xi} = \frac{2}{\pi}\log n + o\big((\log n)^{2/3}\log\log n\big)$$

with probability at least $1-o\big(\frac{1}{\sqrt{\log\log n}}\big)$.

The argument of Erdös and Offord is combinatorial and very delicate, even by today's standards. Their main idea is to approximate the number of roots by the number of sign changes in $P_n(x_1),\dots,P_n(x_k)$ where $(x_1,\dots,x_k)$ is a carefully chosen deterministic sequence of points of length $k=(\frac{2}{\pi}+o(1))\log n$. The authors showed that with high probability, almost every interval (x_i,x_{i+1}) contains exactly one root, and used this fact to prove Theorem 4.2.

Our main result in this section is the following universality statement.

THEOREM 4.3. (Universality for Kac polynomials) *Let k,l be nonnegative integers with $k+l \geq 1$. Assume that $\xi_0,\dots,\xi_n$ and $\tilde{\xi}_0,\dots,\tilde{\xi}_n$ are real random variables with mean 0, satisfying Condition C1 and the polynomials P_n, $\tilde{P}_n$ are Kac polynomials with respect to these variables. Then there exist positive constants C',c depending only on k,l and the constants in Condition C1 such that the following holds.*

For every $0<\theta_n<1$, for any real numbers $x_1,\dots,x_k$, and complex numbers $z_1,\dots,z_l$ with $1-2\theta_n \leq |x_i|,|z_j| \leq 1-\theta_n+1/n$ for all i,j, and for any function $G: \mathbb{R}^k\times\mathbb{C}^l \to \mathbb{C}$ supported on $\prod_{i=1}^k [x_i-10^{-3}\theta_n, x_i+10^{-3}\theta_n]\times\prod_{j=1}^l B(z_j,10^{-3}\theta_n)$ with continuous derivatives up to order $2(k+l)+4$ and $\|\nabla^a G\|_\infty \leq (\theta_n+1/n)^{-a}$ for all $0\leq a\leq 2(k+l)+4$, we have

$$\left|\mathbf{E}\sum G(\zeta_{i_1},\dots,\zeta_{i_k},\zeta_{j_1},\dots,\zeta_{j_l}) - \mathbf{E}\sum G(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k},\tilde{\zeta}_{j_1},\dots,\tilde{\zeta}_{j_l})\right| \\ \leq C'\theta_n^c + C'n^{-c},$$

where the first sum runs over all $(k+l)$-tuples $(\zeta_{i_1},\dots,\zeta_{i_k},\zeta_{j_1},\dots,\zeta_{j_l})\in\mathbb{R}^k\times\mathbb{C}_+^l$ of the roots $\zeta_1,\zeta_2,\dots$ of P_n, and the second sum runs over all $(k+l)$-tuples $(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k},\tilde{\zeta}_{j_1},\dots,\tilde{\zeta}_{j_l})\in\mathbb{R}^k\times\mathbb{C}_+^l$ of the roots $\tilde{\zeta}_1,\tilde{\zeta}_2,\dots$ of $\tilde{P}_n$.

Remark 4.4. Theorem 4.3 provides universality result for the polynomial P_n on the disk $B(0,1+1/n)$. For the complement of this disk, consider $Q_n(z) := z^n P_n(z^{-1})$ which is another Kac polynomial. Since the roots of Q_n are just the reciprocal of the roots of P_n, the universality of Q_n in $B(0,1)$ implies the universality of P_n outside the disk $B(0,1)$.

As a corollary, we get the following result on the number of real roots of these polynomials which recovers the main result of Do and the authors in [11].

COROLLARY 4.5. *Let C be a positive constant. Assume that the random variables ξ_i are independent (not necessarily identically distributed) real random variables with mean 0, variance 1 and $(2+\varepsilon)$-moments bounded by C. Then*

$$\mathbf{E}N_{P_n}(\mathbb{R}) = \frac{2}{\pi}\log n + O(1)$$

where the implicit constant depends only on C and ε.

Theorem 4.3 strengthens an earlier result of Tao and the second author [58]. The result in [58] only covers the bulk of the spectrum, namely the region $1-n^{-\varepsilon}\leq |x| \leq 1+n^{-\varepsilon}$. Restricting to the number of real roots, it yields

$$\mathbf{E}N_{P_n}(\mathbb{R}) = O(\log n)$$

instead of the more precise (and optimal) estimate in Corollary 4.5. Another new feature is that our result also yields sharp estimates for the size of level sets $\{z\in\mathbb{C}:$

$P_n(z) = a\}$, for any fixed a, since we can allow that in Theorem 4.3 and Corollary 4.5, ξ_0 (and in fact any finite number of ξ_i) has non-zero, bounded mean. Our proofs work automatically under this extension. A version of Corollary 4.5 is obtained earlier in [41] using a different approach that combines the local universality in the bulk and a comparison of the number of real roots of P_n with that of $P_{n'}$ for n' much larger than n.

The proof in [58] made use of a deep anti-concentration lemma [58, Lemma 14.1] whose proof relies on the Inverse Littlewood-Offord theory and a weak quantitative version of Gromov's theorem. The proof we will provide here is simple and almost identical to the one used to treat random trigonometric polynomials in the last section. For random variables having continuous distributions (such as the cases treated by Kac and Stevens mentioned above), the anti-concentration property (see Remark 2.3) is immediate.

Remark 4.6. One can routinely modify the proofs of Theorem 4.3 and Corollary 4.5 to show that these results hold for more general settings. For example, the proofs can be used to show that these results apply for

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

where ξ_i are independent (not necessarily identically distributed) random variables satisfying Condition C1 with zero mean and the deterministic coefficients c_i grow polynomially. Specifically, these results hold for derivatives of the Kac polynomials of any given order. We leave the details to the interested reader. The aforementioned results for this general version were proven in the previous work [11] using much more involved tools and arguments.

We defer the proofs of Theorem 4.3 and Corollary 4.5 to Section 11.

5. Application: Universality for Weyl series. In this section, we discuss an application of our main theorems to Weyl series

$$P(z) = \sum_{j=0}^{\infty} \frac{\xi_j z^j}{\sqrt{j!}}$$

where ξ_j are independent complex random variables satisfying the matching condition C1 with the $\tilde{\xi}_j$ being standard complex Gaussian random variables with density $\frac{1}{\pi} e^{-|z|^2}$. In the literature, Weyl series are also referred to as flat series.

The flat series $\tilde{P}(z) = \sum_{j=0}^{\infty} \frac{\tilde{\xi}_j z^j}{\sqrt{j!}}$ is also known as the flat Gaussian analytic function and has been studied intensively over the past few decades. See, for example, [26, 51, 52], and the references therein. Using the Edelman-Kostlan formula [14], one can show that for any Borel set $B \subset \mathbb{C}$, the expected number of roots of

$\tilde{P}$ in B is

$$\mathbf{E}N_{\tilde{P}}(B) = \frac{1}{\pi}m(B) \tag{7}$$

where $m(B)$ is the Lebesgue measure of B.

For general random variables, to compare the distribution of the roots of P with that of $\tilde{P}$, Kabluchko and Zaporozhets (2014) [30] showed that with probability 1, the rescaled empirical measure μ_r defined by

$$\mu_r(A) = \frac{1}{r}\sum_{\zeta:P(\zeta)=0} \mathbf{1}_{\zeta\in\sqrt{r}A}$$

converges vaguely as $r \to \infty$ to the measure $\frac{1}{\pi}m(\cdot)$, which is, as mentioned above, the corresponding measure for $\tilde{P}$. We recall that a sequence of measures (μ_r) is said to converge vaguely to a measure μ if $\lim_{r\to\infty}\int f d\mu_r = \int f d\mu$ for every continuous, compactly supported function f.

The aforementioned result of [30] is about the rescaled measures μ_r. Thus, it provides an asymptotically sharp estimate on the number of roots of P in large domains of the form $\sqrt{r}B$ where $r \to \infty$ and B is a fixed “nice” measurable domain, but does not give estimates for the number of roots in domains with fixed area, as in (7).

Using our framework, we obtain the following result at the local scale.

THEOREM 5.1. (Universality for random flat series) *Assume that the complex random variables ξ_j satisfy the matching condition C1 with the $\tilde{\xi}_j$ being standard complex Gaussian random variables and the random variables* $\mathrm{Re}(\xi_0), \mathrm{Im}(\xi_0), \mathrm{Re}(\xi_1), \mathrm{Im}(\xi_1), \dots$ *are independent. Then there exist positive constants C, c depending only on the constants in Condition C1 such that the following holds.*

For any complex number z_0 and for any function $G : \mathbb{C} \to \mathbb{C}$ supported on $B(z_0, 1)$ with continuous derivatives up to order 6 and $\|\nabla^a G\|_\infty \le 1$ for all $0 \le a \le 6$, we have

$$\left|\mathbf{E}\sum G(\zeta) - \mathbf{E}\sum G(\tilde{\zeta})\right| \le C|z_0|^{-c},$$

where the first sum runs over all the roots $\zeta_1, \zeta_2, \dots$ of P, and the second sum runs over all the roots $\tilde{\zeta}_1, \tilde{\zeta}_2, \dots$ of $\tilde{P}$.

As a corollary, we obtain a sharp estimate on the number of roots in regions with a fixed area.

COROLLARY 5.2. *For any constant $C > 0$, let B be an angular square $B = \{Re^{i\theta} : R \in [r, r+1], \theta \in [\theta_0, \theta_0 + C/r]\}$ for some numbers $r > 0$ and θ_0. Under the assumption of Theorem 5.1, we have*

$$\mathbf{E}N_P(B) = \frac{1}{\pi}m(B) + O(r^{-c}) \quad \text{as } r \to \infty,$$

where c and the implicit constant only depend on C and the constants in Condition C1.

The angular square B can be replaced by a disk, square, or any other nice domains whose indicator functions can be well approximated by smooth functions, with only a nominal modification of the proof. Thus, we have a generalization of (7) for flat series with general random coefficients.

To the best of our knowledge, Theorem 5.1 and Corollary 5.2 are new. We present a short proof of these results in Section 12.

6. Application: Universality for elliptic polynomials. In this section, we briefly illustrate how to apply our framework to elliptic polynomials

$$P_n(z) = \sum_{i=0}^{n} \sqrt{\binom{n}{i}} \xi_i z^i.$$

where ξ_j are independent real random variables satisfying the matching condition C1 with the $\tilde{\xi}_j$ being standard real Gaussian random variables.

For the Gaussian case, the polynomial $\tilde{P}_n(z) = \sum_{i=0}^{n} \sqrt{\binom{n}{i}} \tilde{\xi}_i z^i$ has exactly $\sqrt{n}$ real roots in expectation (see, for example, [5, 14]). In their paper [6], among other results, Bleher and Di extended this result to the non-Gaussian setting.

THEOREM 6.1. [6, Theorem 5.3] *Let ξ_j be iid random variables with mean* 0 *and variance* 1. *Assume furthermore that they are continuously distributed with sufficiently smooth density. Then*

$$\lim_{n\to\infty} \frac{\mathbf{E} N_{P_n}(\mathbb{R})}{\sqrt{n}} = 1.$$

We refer the reader to the original paper [6] for the precise description of "sufficiently smooth". The same result with this assumption being removed is obtained in a recent work of Flasche-Kabluchko [21].

Later, Tao and the second author in [58, Theorem 5.6] showed that the same result holds when the random variables ξ_j are only required to be independent with mean 0, variance 1, and finite $(2+\varepsilon)$-moments. Here we apply our framework to recover these results assuming the more flexible Condition C1, which allows a constant number of ξ_j to have non-zero means. Let us first start with a local universality result.

THEOREM 6.2. (Universality for random elliptic polynomials) *Assume that the real random variables ξ_j are independent and satisfy the matching condition C1 with the $\tilde{\xi}_j$ being standard real Gaussian random variables. Then there exist positive constants C, c depending only on the constants in Condition C1 such that the following holds.*

For any real number x_0 with $n^{-1/2+\varepsilon} \leq |x_0| \leq 1$ and for any function $G : \mathbb{C} \to \mathbb{C}$ supported on $[x_0 - 1/\sqrt{n}, x_0 + 1/\sqrt{n}]$ with continuous derivatives up to order 6 and $\|\nabla^a G\|_\infty \leq n^{a/2}$ for all $0 \leq a \leq 6$, we have

$$\left|\mathbf{E}\sum G(\zeta) - \mathbf{E}\sum G(\tilde{\zeta})\right| \leq Cn^{-c},$$

where the first sum runs over all roots $\zeta_1, \zeta_2, \ldots$ of P_n, and the second sum runs over all the roots $\tilde{\zeta}_1, \tilde{\zeta}_2, \ldots$ of $\tilde{P}_n$.

Remark 6.3. If P_n satisfies the assumptions of Theorem 6.2, so does the polynomial $Q_n(z) = z^n P_n\left(\frac{1}{z}\right) = \sum_{i=0}^{n} \sqrt{\binom{n}{i}} \xi_{n-i} z^i$. And since the roots of Q_n are just the reciprocals of the roots of P_n, from the conclusion of Theorem 6.2 for Q_n, one can obtain the corresponding universality result of P_n on the domain $1 \leq |x_0| \leq n^{1/2-\varepsilon}$.

Thanks to this remark, our result proves universality on the domain $n^{-1/2+\varepsilon} \leq |x_0| \leq n^{1/2-\varepsilon}$. By showing that the contribution outside of this domain is negligible, we obtain the following more quantitative version of Theorem 6.1.

COROLLARY 6.4. *Under the assumption of Theorem 6.2, we have*

$$\mathbf{E}N_{P_n}(\mathbb{R}) = \sqrt{n} + O\left(n^{1/2-c}\right)$$

where c and the implicit constant only depend on the constants in Condition C1.

We give a short proof of these results in Section 13. We note that a corresponding statement can be made concerning the expected number of real roots on a fixed interval $[a, b] \subset \mathbb{R}$, using the same proof.

7. Application: Universality for random Taylor series. Let Γ denote the Gamma function. In a recent paper [20], Flasche and Kabluchko considered the following random series

$$P(x) = \sum_{k=0}^{\infty} \xi_k c_k x^k$$

where the c_k are real deterministic coefficients such that

$$c_k^2 = \frac{k^{\gamma-1}}{\Gamma(\gamma)} L(k)$$

for some constant $\gamma > 0$ and some function $L : (0, \infty) \to \mathbb{R}$ satisfying $L(t) > 0$ for sufficiently large t and $\lim_{t\to\infty} \frac{L(\lambda t)}{L(t)} = 1$ for all $\lambda > 0$. For example, $L(x)$ is some power of $\log x$.

We follow the terminology in [20] and call such a function L a *slowly varying* function and the function P a *random series with regularly varying coefficients*. The following is the main result of [20].

THEOREM 7.1. [20, Theorem 1.1] *Assume that the random variables ξ_k are iid real random variables with zero mean and unit variance. Then*

$$\lim_{r\uparrow 1}\frac{\mathbf{E}N_P[0,r]}{-\log(1-r)}=\frac{\sqrt{\gamma}}{2\pi}.$$

We reprove Theorem 7.1 under the (slightly different) assumption that the random variables ξ_k are independent (not necessarily identically distributed) real random variables with zero mean, unit variance, and uniformly bounded $(2+\varepsilon)$-moments. As usual, we allow that a few random variables have non-zero bounded mean, and so our result also applies to level sets. Our method also yields a polynomial rate of convergence.

As before, we obtain this as a corollary of a stronger theorem establishing the local universality of the roots. Let

$$\tilde{P}(x)=\sum_{k=0}^{\infty}\tilde{\xi}_k c_k x^k$$

where the $\tilde{\xi}_k$ are independent standard Gaussian.

THEOREM 7.2. (Universality for random series with regularly varying coefficients) *Let k,l be nonnegative integers with $k+l\geq 1$. Assume that the real random variables ξ_j are independent and satisfy the matching condition C1 with the $\tilde{\xi}_j$ being standard real Gaussian random variables. There exist positive constants C',c depending only on the constants in Condition C1 such that the following holds.*

Let $0<\delta<1$, and let $x_1,\dots,x_k$ be real numbers and $z_1,\dots,z_l$ be complex numbers satisfying $1-2\delta\leq|x_i|,|z_j|\leq 1-\delta$ for all relevant i,j. Let $G:\mathbb{R}^k\times\mathbb{C}^l\to\mathbb{C}$ by a function supported on $\prod_{i=1}^k[x_i-10^{-3}\delta,x_i+10^{-3}\delta]\times\prod_{j=1}^l B(z_j,10^{-3}\delta)$ with continuous derivatives up to order $2(k+l)+4$ and $\|\nabla^a G\|_\infty\leq\delta^{-a}$ for all $0\leq a\leq 2(k+l)+4$. Then

$$\left|\mathbf{E}\sum G(\zeta_{i_1},\dots,\zeta_{i_k},\zeta_{j_1},\dots,\zeta_{j_l})-\mathbf{E}\sum G(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k},\tilde{\zeta}_{j_1},\dots,\tilde{\zeta}_{j_l})\right|\leq C'\delta^c,$$

where the first sum runs over all $(k+l)$-tuples $(\zeta_{i_1},\dots,\zeta_{i_k},\zeta_{j_1},\dots,\zeta_{j_l})\in\mathbb{R}^k\times\mathbb{C}_+^l$ of the roots $\zeta_1,\zeta_2,\dots$ of P, and the second sum runs over all $(k+l)$-tuples $(\tilde{\zeta}_{i_1},\dots,\tilde{\zeta}_{i_k},\tilde{\zeta}_{j_1},\dots,\tilde{\zeta}_{j_l})\in\mathbb{R}^k\times\mathbb{C}_+^l$ of the roots $\tilde{\zeta}_1,\tilde{\zeta}_2,\dots$ of $\tilde{P}$.

COROLLARY 7.3. *Under the assumption of Theorem 7.2, there exist positive constants C' and c such that the following hold.*

(1) *For any* $r \in (0,1)$,

$$\left|\mathbf{E}N_P[0,r] - \mathbf{E}N_{\tilde{P}}[0,r]\right| \le C$$

where $N_P[0,r]$ *and* $N_{\tilde{P}}[0,r]$ *are the number of real roots of* P *and* $\tilde{P}$ *in* $[0,r]$*, respectively.*

(2) *We have*

$$\lim_{r\uparrow 1} \frac{\mathbf{E}N_P[0,r]}{-\log(1-r)} = \frac{\sqrt{\gamma}}{2\pi}.$$

We prove Theorem 7.2 and Corollary 7.3 in Section 14.

After this paper has been finished, the authors become aware of a very recent and interesting result of Flasche-Kabluchko [21] in which a completely different method is developed to study systematically the elliptic polynomial, Weyl polynomial, flat random analytic function, and hyperbolic random analytic function. As Flasche and Kabluchko mentioned in their paper, a similar approach has been applied to random trigonometric polynomials [19] and random Taylor series [20]. Here we draw a quick comparison of the results.

• The results in [21] prove the universality of the density functions, while our results prove universality of all correlation functions. The authors of [21] do not seem to be aware of our paper (which was put on the arxiv several months earlier) and made a comparison with [58]. However, the main result of [58] is also about universality of all correlation functions, but this critical point has been ignored.

• [21] and related papers require that the random variables are identically distributed with finite second moment; our method requires $(2+\varepsilon)$-moment, but the variables do not need to be iid.

• The results in [19, 20, 21] provide the limits as $n \to \infty$. Our results prove the limits with quantitative error terms.

• Our method allows the coefficients to fluctuate. Specifically, in most of the applications in the above sections, a result stated for a random function

$$F(x) = \sum_k \xi_k \phi_k(z)$$

can readily be generalized (with no significant changes in the proofs) to a random function

$$G(x) = \sum_k c_k \xi_k \phi_k(z),$$

where c_k are deterministic coefficients that can take any values in the interval $[1/2, 2]$ (say). In this respect, the method in [21] which relies on assumptions such as [21, Equation (6)] may be more susceptible to coefficients' fluctuations.

8. Proof of Theorems 2.5 and 2.6. Before starting the proofs, let us mention two Jensen's inequalities that we use several times in this manuscript. It will be clear in the context which Jensen's inequality is used. The first, and perhaps more popular, Jensen's inequality relates the value of a convex function of an integral to the integral of that convex function. In particular, for any convex function ϕ on the real line and any real integrable random variable X, we have

$$\phi\big(\mathbf{E}(X)\big) \le \mathbf{E}\phi(X).$$

The second Jensen's inequality provides an upper bound on the number of roots of an analytic function. Assume that f is an analytic function on an open domain that contains the closed disk $\bar{B}(z,R)$. Then for any $r < R$, we have

$$N\big(B(z,r)\big) \le \frac{\log \frac{M}{m}}{\log \frac{R^2+r^2}{2Rr}} \tag{8}$$

where $N(B(z,r))$ is the number of roots (including multiplicities) of f in the open disk $B(z,r)$ and $M = \max_{w\in\bar{B}(z,R)} |f(w)|$, $m = \max_{w\in\bar{B}(z,r)} |f(w)|$. For completeness, we include a short proof of this inequality in Appendix 15.5.

8.1. Proof of Theorem 2.5. We first state a few lemmas. The first lemma reduces the theorem to the case when the function G *splits*, namely G is a product of functions of a single variable. In many applications, G automatically takes this form. This lemma was proved in [58]. We include a short proof in Appendix 15.1.

LEMMA 8.1. *If Theorem 2.5 holds for every function G of the form*

$$G(w_1,\dots,w_m) = G_1(w_1)\cdots G_k(w_k) \tag{9}$$

where for each $1 \le i \le k$, $G_i : \mathbb{C} \to \mathbb{C}$ *is a function supported in* $B(z_i, 1/50)$ *with continuous derivatives up to order* 3 *and* $\|\nabla^a G_i\|_\infty \le 1$ *for all* $0 \le a \le 3$, *then it holds for any function G satisfying the hypothesis of Theorem 2.5. Similarly for Theorem 2.6.*

The next lemma plays a critical role in our approach, as it shows that the singularity problem at 0 (see the discussion in the last subsection of Section 2) can be dealt with assuming anti-concentration at a single point.

LEMMA 8.2. *Let* $0 < \delta_n, c_2 < 1$ *and let* F_n *be an entire function with* $|F_n(w)| \ge \exp(-\delta_n^{-c_2})$ *for some complex number* w *and* $|F_n(z)| \le \exp(\delta_n^{-c_2})$ *for all* $z \in B(w,3/2)$. *Then*

$$\int_{B(w,1/2)} \big|\log|F_n(z)|\big|^2 dz \le 720^2 \times \delta_n^{-6c_2}.$$

The constant $720^2 = 518400$ is for explicitness and plays no specific role. Both this and the constant 6 in the exponent can be reduced but we make no attempt to optimize these constants. The proof follows from ideas in [11] and is included in Appendix 15.2.

The following lemma shows that the logarithm function satisfies a universality property. It is a variant of a lemma in [58] and we include the proof in Appendix 15.3.

LEMMA 8.3. (Log-comparability) *Assume that the coefficients ξ_i and $\tilde{\xi}_i$ satisfy Condition C1 for some constants N_0, ε, τ. Let α_1 be a positive constant and k be a positive integer. Assume that there exists a constant $C > 0$ such that the random functions F_n and $\tilde{F}_n$ satisfy Condition C2(4) with parameters α_1 and C. There exist positive constants α_0 and C' such that for any $z_1, \dots, z_k \in D_n + B(0, 1/10)$, and function $K : \mathbb{C}^k \to \mathbb{C}$ with continuous derivatives up to order 3 and $\|\nabla^a K\|_\infty \le \delta_n^{-\alpha_0}$ for all $0 \le a \le 3$, we have*

$$\left|\mathbf{E}K\big(\log|F_n(z_1)|, \dots, \log|F_n(z_k)|\big) - \mathbf{E}K\big(\log|\tilde{F}_n(z_1)|, \dots, \log|\tilde{F}_n(z_k)|\big)\right| \le C'\delta_n^{\alpha_0}.$$

Remark 8.4. Following the proof, one can set $\alpha_0 = \frac{3\alpha_1\varepsilon}{10^3}$.

Proof of Theorem 2.5. By Lemma 8.1, we can assume that the function G has the form (9). We need to show that

$$\left|\mathbf{E}\prod_{j=1}^k\left(\sum_i G_j(\zeta_i)\right) - \mathbf{E}\prod_{j=1}^k\left(\sum_i G_j(\tilde{\zeta}_i)\right)\right| \le C'\delta_n^c, \tag{10}$$

for some constant $c > 0$. By Green's formula, we have

$$\sum_i G_j(\zeta_i) = \int_{\mathbb{C}} \log|F_n(z)| H_j(z)dz = \int_{B(z_j, 1/10)} \log|F_n(u_j)| H_j(u_j)du_j, \tag{11}$$

where $H_j(z) = \frac{1}{2\pi}\triangle G_j(z)$. Note that $\mathrm{supp}(H_j) \subset B(z_j, 1/10)$ and $\|H_j\|_\infty \le 1$ for all $z \in \mathbb{C}$, thanks to the assumption on G in Theorem 2.5. (As usual, $\|f\|_\infty = \sup_{z\in\mathbb{C}}|f(z)|$.) When F_n is identically 0, we assume by convention that the left-hand side and the right-hand side are 0.

Let A be a sufficiently large constant and c_1 be a sufficiently small positive constant. For this proof, it suffices to set $c_1 := \frac{\alpha_0}{300k^2}$ and $A := 2kC_1 + \frac{\alpha_1\varepsilon}{60}$. This choice, together with the value of α_0 in Remark 8.4, yields the explicit values of A and c_1 in the theorem.

Let $\bar{c}_1 := 100kc_1$. The power c in (10) can be chosen (quite generously) to be c_1.

Let $K : \mathbb{R} \to \mathbb{R}$ be a smooth function with the following properties:

- K is supported on the interval $[-2\delta_n^{-\bar{c}_1}, 2\delta_n^{-\bar{c}_1}]$.

- $K(x) = x$ for all $x \in [-\delta_n^{-\bar{c}_1}, \delta_n^{-\bar{c}_1}]$.
- $\|K^{(a)}\|_\infty = O(\delta_n^{-\bar{c}_1})$ for all $0 \le a \le 3$ (where $K^{(l)}$ is the lth derivative of K).
- $|K(x)| \le |x|$ for all $x \in \mathbb{R}$.

Let $\Gamma := \prod_{j=1}^k B(z_j, 1/10)$ and $H(u) := \prod_{j=1}^k H_j(u_j)$ for $u := (u_1, \dots, u_k)$. By (11), we have

$$\mathbf{E}\prod_{j=1}^{k}\left(\sum_i G_j(\zeta_i)\right) = \mathbf{E}\int_\Gamma H(u)\prod_{j=1}^{k}\log|F_n(u_j)|du = A_1 + A_2$$

where

$$A_1 := \mathbf{E}\int_\Gamma H(u)\prod_{j=1}^{k}K\big(\log|F_n(u_j)|\big)du,$$

$$A_2 := \mathbf{E}\int_\Gamma H(u)\left[\prod_{j=1}^{k}\log|F_n(u_j)| - \prod_{j=1}^{k}K\big(\log|F_n(u_j)|\big)\right]du.$$

Let $\tilde{A}_1$ and $\tilde{A}_2$ be the corresponding terms for $\tilde{F}_n$. Our goal is to show that

$$A_1 + A_2 - \tilde{A}_1 - \tilde{A}_2 = O\big(\delta_n^c\big). \tag{12}$$

By Lemma 8.3, we have $A_1 - \tilde{A}_1 = O(\delta_n^{\bar{c}_1})$. We next show that both A_2 and $\tilde{A}_2$ are of order $O(\delta_n^{c_1})$. It suffices to consider A_2, as the treatment of $\tilde{A}_2$ is similar.

Let $\mathcal{A}_0$ be the event on which the following two properties hold.

- For all $1 \le j \le k$, $|F_n(z_j')| \ge \exp(-\delta_n^{-c_1})$ for some $z_j' \in B(z_j, 1/100)$.
- $|F_n(z)| \le \exp(\delta_n^{-c_1})$ for all $z \in B(z_j, 2)$.

By Conditions C2(2) and C2(3), $\mathbf{P}(\mathcal{A}_0^c) \le C\delta_n^A$, where $\mathcal{A}_0^c$ is the complement of $\mathcal{A}_0$. We next break up A_2 as follows

$$\begin{aligned} A_2 &= \mathbf{E}\int_\Gamma H(u)\left[\prod_{j=1}^{k}\log|F_n(u_j)| - \prod_{j=1}^{k}K\big(\log|F_n(u_j)|\big)\right]du\mathbf{1}_{\mathcal{A}_0} \\ &\quad + \mathbf{E}\int_\Gamma H(u)\prod_{j=1}^{k}\log|F_n(u_j)|du\mathbf{1}_{\mathcal{A}_0^c} - \mathbf{E}\int_\Gamma H(u)\prod_{j=1}^{k}K\big(\log|F_n(u_j)|\big)du\mathbf{1}_{\mathcal{A}_0^c} \\ &=: A_3 + A_4 - A_5. \end{aligned}$$

For A_5, since $\|K\|_\infty \le 2\delta_n^{-\bar{c}_1}$ by construction and $A \ge 2k\bar{c}_1$, we have

$$|A_5| \le 2\delta_n^{-k\bar{c}_1}\mathbf{P}\big(\mathcal{A}_0^c\big) \le 2C\delta_n^{A-k\bar{c}_1} = O\big(\delta_n^{\bar{c}_1}\big) = O\big(\delta_n^{c_1}\big).$$

To bound A_4, from (11) and the boundedness of H_j, we have

$$\left|\int_{B(z_j,1/100)}\log|F_n(u_j)|H_j(u_j)du_j\right| \le N_{F_n}\big(B(z_j,1/100)\big) =: N_j.$$

By Hölder's inequality for products,

$$|A_4| \le \prod_{j=1}^{k} \big(\mathbf{E} N_j^k \mathbf{1}_{\mathcal{A}_0^c}\big)^{1/k}.$$

We bound each term on the right using Hölder's inequality as follows

$$\mathbf{E} N_j^k \mathbf{1}_{\mathcal{A}_0^c} \le \delta_n^{-kC_1} \mathbf{P}\big(\mathcal{A}_0^c\big) + \big(\mathbf{E} N_j^{k+1} \mathbf{1}_{N_j \ge \delta_n^{-C_1}}\big)^{k/(k+1)} \big(\mathbf{P}\big(\mathcal{A}_0^c\big)\big)^{1/(k+1)}.$$

In our setting, $A \ge kC_1 + (k+1)\bar{c}_1$, the first term on the right-hand side is $O(\delta_n^{c_1})$. Moreover, Condition C2(1) implies that the second term is $O(\mathbf{P}(\mathcal{A}_0^c)^{1/(k+1)}) = O(\delta_n^{c_1})$. Thus, $A_4 = O(\delta_n^{c_1})$.

Finally, to bound A_3, we let B be the (random) set of all $u \in \Gamma$ on which $|\log|F_n(u_j)|| \ge \delta_n^{-\bar{c}_1}$ for some j. Notice that if $u = (u_1, \dots, u_k) \notin B$, then $K(\log|F_n(u_j)|) = \log|F_n(u_j)|$ by the properties of K and the definition of B. Moreover, for $u \in B$, $|K(\log|F_n(u_j)|)| \le |\log|F_n(u_j)||$ as $|K(x)| \le |x|$ for all x. It follows that

$$|A_3| \le 2\mathbf{E} \int_\Gamma \left| \prod_{j=1}^{k} \log\big|F_n(u_j)\big| \right| \mathbf{1}_B(u) du \mathbf{1}_{\mathcal{A}_0}. \tag{13}$$

By Hölder's inequality, the right-hand side is at most

$$2 \left[\mathbf{E} \int_\Gamma \left| \prod_{j=1}^{k} \log|F_n(u_j)| \right|^2 du \mathbf{1}_{\mathcal{A}_0} \right]^{1/2} \left[\mathbf{E} \int_\Gamma \mathbf{1}_B(u) du \mathbf{1}_{\mathcal{A}_0} \right]^{1/2}.$$

By Lemma 8.2, on the event $\mathcal{A}_0$, we have

$$\int_{B(z_j,1/100)} \big| \log|F_n(u_j)| \big|^2 du_j = O\big(\delta_n^{-6c_1}\big). \tag{14}$$

It follows that

$$\int_\Gamma \left| \prod_{j=1}^{k} \log\big|F_n(u_j)\big| \right|^2 du = O\big(\delta_n^{-6kc_1}\big).$$

On the other hand, by the definition of B,

$$\int_\Gamma \mathbf{1}_B(u) du \mathbf{1}_{\mathcal{A}_0} = O\left(\mathbf{1}_{\mathcal{A}_0} \sum_{j=1}^{k} \int_{B(z_j,1/100)} \mathbf{1}_{|\log|F_n(u_j)|| \ge \delta_n^{-\bar{c}_1}} du_j \right).$$

Furthermore,

$$\int_{B(z_j,1/100)} \mathbf{1}_{|\log|F_n(u_j)|| \ge \delta_n^{-\bar{c}_1}} du_j \le \delta_n^{2\bar{c}_1} \int_{B(z_j,1/100)} \big| \log|F_n(z)| \big|^2 dz.$$

Using (14), we obtain

$$\mathbf{E}\int_\Gamma \mathbf{1}_B(u)du\mathbf{1}_{\mathcal{A}_0} = O\big(\delta_n^{2\bar{c}_1}\delta_n^{-6kc_1}\big).$$

It follows that

$$|A_3| = O\Big(\big(\delta_n^{-6kc_1}\times\delta_n^{2\bar{c}_1}\delta_n^{-6kc_1}\big)^{1/2}\Big) = O\big(\delta_n^{\bar{c}_1-6kc_1}\big) = O\big(\delta_n^{c_1}\big)$$

as we set $\bar{c}_1 > 7kc_1$. The bounds on $|A_3|, |A_4|$ and $|A_5|$ together imply $|A_2| = O(\delta_n^{c_1})$, concluding the proof. □

8.2. Proof of Theorem 2.6. By Lemma 8.1, it suffices to assume that G can be decomposed into functions of single variables, namely

$$G(x_1,\dots,x_k,z_1,\dots,z_l) = H_1(x_1)\dots H_k(x_k)G_1(z_1)\dots G_l(z_l)$$

where the $H_i:\mathbb{R}\to\mathbb{C}$ and $G_j:\mathbb{C}\to\mathbb{C}$ are smooth functions supported on $[x_i - 1/50, x_i+1/50]$ and $B(z_j,1/50)$ (respectively) and satisfying

$$|\nabla^a H_i(x)|, |\nabla^a G_j(z)| \le 1$$

for any $x\in\mathbb{R}$, $z\in\mathbb{C}$ and $0\le a\le 3$.

In other words, one needs to show that

$$\left|\mathbf{E}\left(\prod_{i=1}^k X_i\right)\left(\prod_{j=1}^l Y_j\right) - \mathbf{E}\left(\prod_{i=1}^k \tilde{X}_i\right)\left(\prod_{j=1}^l \tilde{Y}_j\right)\right| \le C'\delta_n^{\bar{c}}, \tag{15}$$

for some constants $C',\bar{c}>0$, where

$$X_i = \sum_{\zeta_s\in\mathbb{R}} H_i(\zeta_s),\quad \tilde{X}_i = \sum_{\tilde{\zeta}_s\in\mathbb{R}} H_i(\tilde{\zeta}_s),\quad Y_j = \sum_{\zeta_s\in\mathbb{C}_+} G_j(\zeta_s),\quad \tilde{Y}_j = \sum_{\tilde{\zeta}_s\in\mathbb{C}_+} G_j(\tilde{\zeta}_s).$$

(We use $\bar{c}$ instead of c to denote the exponent on the right-hand side, since we reserve c for the exponent in Theorem 2.5, which we will use in the proof.)

The proof follows the ideas in [58]. The first step is to show that the number of complex zeros near the real axis is small with high probability. Let c be the constant exponent in Theorem 2.5 corresponding to $k+l$. Following Remark 2.7, we can set $c = \frac{\alpha_1\varepsilon}{10^5(k+l)^2}$.

With this choice of c, we set $c_2 := \frac{c}{100} = \frac{\alpha_1\varepsilon}{10^7(k+l)^2}$ and $\gamma := \delta_n^{c_2}$. Let us also recall that in the statement of this theorem (Theorem 2.6), $c_1 = \frac{\alpha_1\varepsilon}{10^9(k+l)^4}$, which is much smaller than c_2: $c_1 = \frac{c_2}{100(k+l)^2}$.

LEMMA 8.5. *Under the assumptions of Theorem 2.6, we have*

$$\mathbf{P}\big(N_{F_n}B(x,\gamma)\ge 2\big) = O\big(\gamma^{3/2}\big),\quad \textit{for all } x\in\mathbb{R}\cap\big(D_n + B(0,1/50)\big)$$

where the implicit constant depends only on the constants in Conditions C1 and C2 but not on n, δ_n, D_n *and* x.

The power $3/2$ in the above lemma is not critical, we only need something strictly greater than 1.

Assuming this lemma, the rest of the proof is relatively simple. For every $1 \leq i \leq k$, consider the strip $S_i := [x_i - 1/50, x_i + 1/50] \times [-\gamma/4, \gamma/4]$. We can cover S_i by $O(\gamma^{-1})$ disks of the form $B(x, \gamma)$, where $x \in [x_i - 1/50, x_i + 1/50]$. Since F_n has real coefficients, if z is a root of F_n in $S_i \backslash \mathbb{R}$, so is it conjugate $\bar{z}$. Using Lemma 8.5 and the union bound, we obtain

$$\begin{aligned} &\mathbf{P}\big(\text{there is at least 1 (or equivalently 2) root(s) in } S_i \backslash \mathbb{R}\big) \\ &\qquad = O\big(\gamma^{-1}\gamma^{3/2}\big) = O\big(\gamma^{1/2}\big). \end{aligned} \tag{16}$$

Define $\mathfrak{H}_i(z) := H_i(\mathrm{Re}(z))\phi\big(\frac{4\,\mathrm{Im}(z)}{\gamma}\big)$, where $\phi : \mathbb{R} \to [0,1]$ is a smooth function that is supported on $[-1,1]$, with $\phi(0) = 1$ and $\left\|\phi^{(a)}\right\|_\infty = O(1)$ for all $0 \leq a \leq 3$. It is easy to see that $\mathfrak{H}_i$ is a smooth function supported on S_i with $\|\mathfrak{H}_i\|_\infty \leq 1$, and $\left\|\nabla^a \mathfrak{H}_i\right\|_\infty = O(\gamma^{-a})$ for $0 \leq a \leq 3$.

Set $\mathfrak{X}_i := \sum_s \mathfrak{H}_i(\zeta_s)$ and $D_i := \mathfrak{X}_i - X_i$. By the definitions of $\mathfrak{X}_i$ and X_i, $D_i = \sum_{\zeta_s \notin \mathbb{R}} \mathfrak{H}_i(\zeta_s)$. Our general strategy is to use $\mathfrak{X}_i$ to approximate X_i, then apply Theorem 2.5 to $\mathfrak{X}_i$ and finish the proof using a triangle inequality.

From (16), $D_i = 0$ with probability at least $1 - O(\gamma^{1/2})$. Notice that by the definition of D_i and the fact that $\|\mathfrak{H}_i\|_\infty \leq 1$,

$$|D_i| \leq N_{F_n} B(x_i, 1/5). \tag{17}$$

By (17) and Jensen's inequality (8),

$$|D_i| \leq N_{F_n} B(x_i, 1/5) = O\Big(\log \max_{w \in B(x_i, 2)} |F_n(w)| - \log \max_{z \in B(x_i, 1/5)} |F_n(z)|\Big).$$

By Conditions C2(2) and C2(3), with probability at least $1 - O(\delta_n^A)$, there exists $z \in B(x_i, 1/100)$ such that both terms on the right-hand side are of order $O(\delta_n^{-c_1})$. Therefore, with probability at least $1 - O(\delta_n^A)$, we have $|D_i| \leq N_{F_n} B(x_i, 1/5) \leq C' \delta_n^{-c_1}$ for some constant C'. For the rest of this proof, we denote $N_i := N_{F_n} B(x_i, 1/5)$.

Our next step is to bound $\mathbf{E}|D_i|^{k+l}$. To start, we have

$$\mathbf{E}|D_i|^{k+l} \leq \mathbf{E}\big(|D_i|^{k+l} \mathbf{1}_{N_i \leq C'\delta_n^{-c_1}}\big) + \mathbf{E}\big(N_i^{k+l} \mathbf{1}_{N_i > C'\delta_n^{-c_1}}\big). \tag{18}$$

Since $D_i = 0$ with probability at least $1 - O(\gamma^{1/2})$,

$$\mathbf{E}\big(|D_i|^{k+l} \mathbf{1}_{N_i \leq C'\delta_n^{-c_1}}\big) = O\big(\delta_n^{-c_1(k+l)} \gamma^{1/2}\big) = O\big(\delta_n^{-c_1(k+l) + c_2/2}\big) = O\big(\delta_n^{c_1(k+l)^2}\big)$$

because $c_2 \geq 4c_1(k+l)^2$.

For the second term in (18), we further break up the event $N_i > C'\delta_n^{-c_1}$ into two events

$$\Omega_1 := \delta_n^{-C_1} \geq N_i > C'\delta_n^{-c_1} \quad \text{and} \quad \Omega_2 := \delta_n^{-C_1} \leq N_i$$

where C_1 is the constant in the statement of Theorem 2.6. We have

$$\mathbf{E}N_i^{k+l}\mathbf{1}_{\Omega_1} \leq \delta_n^{-C_1(k+l)}\mathbf{P}(\Omega_1) = O\big(\delta_n^{A-C_1(k+l)}\big) = O\big(\delta_n^{c_1(k+l)^2}\big).$$

Moreover, by Hölder's inequality,

$$\begin{aligned}\mathbf{E}N_i^{k+l}\mathbf{1}_{\Omega_2} &\leq \mathbf{P}\big(\Omega_2\big)^{\frac{2}{k+l+2}}\big(\mathbf{E}N_i^{k+l+2}\mathbf{1}_{\Omega_2}\big)^{\frac{k+l}{k+l+2}} \\ &= O\big(\delta_n^{A/(k+l+2)}\big)\big(\mathbf{E}N_i^{k+l+2}\mathbf{1}_{\Omega_2}\big)^{\frac{k+l}{k+l+2}}.\end{aligned}$$

Under the assumption of Theorem 2.6, Condition C2(1) holds for the parameter $k+l$, which provides $\mathbf{E}N_i^{k+l+2}\mathbf{1}_{\Omega_2} = O(1)$. As we set A to be much larger than c_1, it is easy to check that

$$\mathbf{E}N_i^{k+l}\mathbf{1}_{\Omega_2} = O\big(\delta_n^{A/(k+l+2)}\big) = O\big(\delta_n^{c_1(k+l)^2}\big).$$

Thus,

$$\begin{aligned}\mathbf{E}\big(N_{F_n}B(x_i,1/5)\big)^{k+l}\mathbf{1}_{N_{F_n}B(x_i,1/5)\geq C'\delta_n^{-c_1}} &= O\big(\delta_n^{A/(k+l+2)}\big) \\ &= O\big(\delta_n^{c_1(k+l)^2}\big).\end{aligned} \tag{19}$$

Combining all these bounds with (18), we obtain

$$\mathbf{E}|D_i|^{k+l} = O\big(\delta_n^{c_1(k+l)^2}\big).$$

Moreover, from the above bounds, we get

$$\begin{aligned}\mathbf{E}|\mathfrak{X}_i|^{k+l} \leq \mathbf{E}N_i^{k+l} &= \mathbf{E}N_i^{k+l}\mathbf{1}_{N_i\leq C'\delta_n^{-c_1}} + \mathbf{E}N_i^{k+l}\mathbf{1}_{\Omega_1} + \mathbf{E}N_i^{k+l}\mathbf{1}_{\Omega_2} \\ &= O\big(\delta_n^{-c_1(k+l)}\big),\end{aligned}$$

where the main contribution comes from the first term. Similarly, $\mathbf{E}|X_i|^{k+l} = O(\delta_n^{-c_1(k+l)})$.

Next, for each $1 \leq j \leq l$, let $\mathfrak{G}_j(z) := G_j(z)\varphi(\text{Im}(z)/\gamma)$ where φ is a smooth function on $\mathbb{R}$ supported on $[1/2,\infty)$ with $\varphi = 1$ on $[1,\infty)$ and $\left\|\varphi_\infty^{(a)}\right\| = O(1)$ for all $0 \leq a \leq 3$.

Set $\mathfrak{Y}_j := \sum_s \mathfrak{G}_j(\zeta_s)$. By similar reasoning, we have $\mathbf{E}|\mathfrak{Y}_j - Y_j|^{k+l} = O(\delta_n^{c_1(k+l)^2})$ and

$$\max\big\{\mathbf{E}|\mathfrak{Y}_j|^{k+l}, \mathbf{E}|Y_j|^{k+l}\big\} = O\big(\delta_n^{-c_1(k+l)}\big).$$

Now, we show that the difference

$$\mathbf{E}\left|\left(\prod_{i=1}^{k}X_i\right)\left(\prod_{j=1}^{l}Y_j\right)-\left(\prod_{i=1}^{k}\mathfrak{X}_i\right)\left(\prod_{j=1}^{l}\mathfrak{Y}_j\right)\right|$$

is small. Using the "telescopic sum" argument, we decompose the difference inside the absolute value sign into the sum of $k+l$ differences, in each of which exactly one of the $X_1,\dots,X_k,Y_1,\dots,Y_j$ is replaced by its counterpart, and then use the triangle inequality to finish. Let us bound the first difference; the argument for the rest is the same. By Hölder's inequality and the previous bounds on D_i,X_i,Y_i etc, we have

$$\begin{aligned}
&\mathbf{E}\left|X_1\left(\prod_{i=2}^{k}X_i\right)\left(\prod_{j=1}^{l}Y_j\right)-\mathfrak{X}_1\left(\prod_{i=2}^{k}X_i\right)\left(\prod_{j=1}^{l}Y_j\right)\right|\\
&\le\left(\mathbf{E}|D_1|^{k+l}\right)^{\frac{1}{k+l}}\prod_{i=2}^{k}\left(\mathbf{E}|X_i|^{k+l}\right)^{\frac{1}{k+l}}\prod_{j=1}^{l}\left(\mathbf{E}|Y_j|^{k+l}\right)^{\frac{1}{k+l}}\\
&=O\left(\delta_n^{c_1(k+l)}\prod_{k+l-1\text{ terms}}\delta_n^{-c_1}\right)=O\left(\delta_n^{c_1}\right).
\end{aligned}$$

Thus,

$$\mathbf{E}\left|\left(\prod_{i=1}^{k}X_i\right)\left(\prod_{j=1}^{l}Y_j\right)-\left(\prod_{i=1}^{k}\mathfrak{X}_i\right)\left(\prod_{j=1}^{l}\mathfrak{Y}_j\right)\right|=O\left(\delta_n^{c_1}\right).$$

We can obtain the same bound for the corresponding terms of $\tilde{F}_n$. Finally, from Theorem 2.5, we have

$$\left|\mathbf{E}\left(\prod_{i=1}^{k}\mathfrak{X}_i\right)\left(\prod_{j=1}^{l}\mathfrak{Y}_j\right)-\mathbf{E}\left(\prod_{i=1}^{k}\tilde{\mathfrak{X}}_i\right)\left(\prod_{j=1}^{l}\tilde{\mathfrak{Y}}_j\right)\right|=O\left(\delta_n^{c_1}\right).$$

The desired estimate now follows from the triangle inequality.

Proof of Lemma 8.5. The first step is to use Theorem 2.5 to reduce to the Gaussian case. Borrowing ideas from [26, Chapter 2], we handle the Gaussian case using Rouché's theorem and various probabilistic estimates based on some properties of the Gaussian distribution.

For this proof, we let $\tilde{\xi}_1,\dots,\tilde{\xi}_n$ be Gaussian random variables with unit variance and satisfying $\mathbf{E}\tilde{\xi}_i=\mathbf{E}\xi_i$ for each $1\le i\le n$.

Let $H:\mathbb{C}\to[0,1]$ be a nonnegative smooth function supported on $B(x,2\gamma)$, such that $H=1$ on $B(x,\gamma)$ and $|\nabla^a H|\le C\gamma^{-a}$ for all $0\le a\le 8$.

Applying Theorem 2.5 to H, we obtain

$$\begin{aligned}\mathbf{P}\big(N_{F_n}B(x,\gamma)\geq 2\big) &\leq \mathbf{E}\sum_{i\neq j}H(\zeta_i)H(\zeta_j)\\ &\leq \mathbf{E}\sum_{i\neq j}H(\tilde{\zeta}_i)H(\tilde{\zeta}_j)+O\big(\delta_n^c\gamma^{-8}\big).\end{aligned}\tag{20}$$

The definition of γ guarantees (via a trivial calculation) that $O(\delta_n^c\gamma^{-8}) = O(\gamma^{3/2})$, with room to spare. Thus, it remains to show

$$\mathbf{E}\sum_{i\neq j}H(\tilde{\zeta}_i)H(\tilde{\zeta}_j)=O\big(\gamma^{3/2}\big).\tag{21}$$

Set $N := N_{\tilde{F}_n}B(x,2\gamma)$; we bound the LHS of (21) from above by

$$\mathbf{E}N^2\mathbf{1}_{N\geq C'\delta_n^{-c_1}}+\mathbf{E}N(N-1)\mathbf{1}_{N<C'\delta_n^{-c_1}}.\tag{22}$$

Using the same argument as in the proof of (19), we can show that

$$\mathbf{E}N^2\mathbf{1}_{N\geq C'\delta_n^{-c_1}}=O\big(\delta_n^{A/(k+l+2)}\big)=O\big(\gamma^{3/2}\big).$$

Thus, it remains to show that $\mathbf{E}N(N-1)\mathbf{1}_{N<C'\delta_n^{-c_1}}=O(\gamma^{3/2})$. Since

$$\mathbf{E}N(N-1)\mathbf{1}_{N<C'\delta_n^{-c_1}}\leq C'^2\delta_n^{-2c_1}\mathbf{P}(N\geq 2),$$

it suffices to prove

$$\mathbf{P}(N\geq 2)=\mathbf{P}\big(N_{\tilde{F}_n}B(x,2\gamma)\geq 2\big)=O\big(\delta_n^{2c_1}\gamma^{3/2}\big).\tag{23}$$

Thus, we have reduced the problem to the Gaussian setting. Let $g(z) := \tilde{F}_n(x)+\tilde{F}_n'(x)(z-x)$ and $p(z) := \tilde{F}_n(z)-g(z)$. By Condition C2(4), for any fixed x, we have $\tilde{F}_n(x)\tilde{F}_n'(x)\neq 0$ with probability 1. So, $g(z)$ has exactly one root. Thus, by Rouché's theorem,

$$\mathbf{P}\big(N_{\tilde{F}_n}B(x,2\gamma)\geq 2\big)\leq \mathbf{P}\left(\min_{z\in\partial B(x,2\gamma)}|g(z)|\leq \max_{z\in\partial B(x,2\gamma)}|p(z)|\right).$$

In the rest of the proof, we bound the right-hand side. We are going to show that with (appropriately) high probability, $\min_{z\in\partial B(x,2\gamma)}|g(z)|$ is not too small and $\max_{z\in\partial B(x,2\gamma)}|p(z)|$ is not too large.

For every $z\in B(x,4\gamma)$, we have $p(z)=\sum_{j=1}^n\tilde{\xi}_jv_j(z)$ where $v_j(z)=\phi_j(z)-\phi_j(x)+(z-x)\phi_j'(z)$. Thus

$$|v_j(z)|\leq |z-x|^2\sup_{w\in B(x,2\gamma)}\big|\phi_j''(w)\big|=O\left(\gamma^2\sup_{w\in B(x,2\gamma)}|\phi_j''(w)|\right).$$

By Condition C2(5),

$$
\begin{aligned}
|\mathbf{E}p(z)| &= O\left(\gamma^2 \sum_{j=1}^n |\mathbf{E}\tilde{\xi}_j| \sup_{w\in B(x,1)} |\phi_j''(z)|\right) \\
&= O\left(\delta_n^{2c_2-c_1} \sqrt{\sum_{j=1}^n |\phi_j(x)|^2}\right),
\end{aligned} \tag{24}
$$

and

$$
\begin{aligned}
\mathrm{Var}(p(z)) &= O\left(\gamma^4 \sum_{i=1}^n \sup_{w\in B(x,2\gamma)} |\phi_j''(w)|^2\right) = O\left(\delta_n^{4c_2-c_1} \sum_{j=1}^n |\phi_j(x)|^2\right) \\
&= O\big(\delta_n^{4c_2-c_1}\,\mathrm{Var}\big(\tilde{F}_n(x)\big)\big).
\end{aligned} \tag{25}
$$

Set $t := \delta_n^{2c_2-c_1}\sqrt{\mathrm{Var}(\tilde{F}_n(x))}$. The previous estimates show that $|\mathbf{E}p(z)| = O(t)$ and $\mathrm{Var}(p(z)) = O(t^2\delta_n^{c_1})$ for all $z \in B(x,4\gamma)$. We will show the following concentration inequality

$$
\begin{aligned}
&\mathbf{P}\left(\max_{z\in\partial B(x,2\gamma)} |p(z) - \mathbf{E}p(z)| \geq \frac{1}{2}t\right) \\
&\quad = O(1)\exp\left(-\frac{t^2}{100\max_{z\in B(x,4\gamma)}\mathrm{Var}(p(z))}\right) = O\big(\gamma^{16/10}\delta_n^{2c_1}\big).
\end{aligned} \tag{26}
$$

Set $\bar{p}(z) := p(z) - \mathbf{E}p(z)$. For any $z \in \partial B(x,2\gamma)$, by Cauchy's integral formula,

$$
\begin{aligned}
|\bar{p}(z)| &\leq \int_0^{2\pi} \frac{|\bar{p}(x+4\gamma e^{i\theta})|}{|z-x-4\gamma e^{i\theta}|} 4\gamma \frac{d\theta}{2\pi} \leq 2\int_0^{2\pi} |\bar{p}(x+4\gamma e^{i\theta})| \frac{d\theta}{2\pi} \\
&\leq \max_{w\in B(x,4\gamma)} \sqrt{\mathrm{Var}(p(w))} \int_0^{2\pi} \frac{|\bar{p}(x+4\gamma e^{i\theta})|}{\sqrt{\mathrm{Var}(\bar{p}(x+4\gamma e^{i\theta}))}} \frac{d\theta}{2\pi}.
\end{aligned}
$$

Hence, by Markov's inequality,

$$
\begin{aligned}
&\mathbf{P}\Big(\max_{z\in\partial B(x,2\gamma)} |\bar{p}(z)| \geq t\Big) \\
&\leq \mathbf{E}\exp\left(\left(\int_0^{2\pi} \frac{\big|\bar{p}(x+4\gamma e^{i\theta})\big|}{10\sqrt{\mathrm{Var}\big(\bar{p}(x+4\gamma e^{i\theta})\big)}} \frac{d\theta}{2\pi}\right)^2\right) e^{-t^2/100\max_{z\in B(x,4\gamma)}\mathrm{Var}(p(z))}.
\end{aligned}
$$

Using Jensen's inequality for convex functions and Fubini's theorem, we obtain

$$\mathbf{E}\exp\left(\left(\int_0^{2\pi}\frac{|\bar{p}(x+4\gamma e^{i\theta})|}{10\sqrt{\mathrm{Var}(\bar{p}(x+4\gamma e^{i\theta}))}}\frac{d\theta}{2\pi}\right)^2\right) \leq \int_0^{2\pi}\mathbf{E}\exp\left(\frac{|\bar{p}(x+4\gamma e^{i\theta})|^2}{100\mathrm{Var}(\bar{p}(x+4\gamma e^{i\theta}))}\right)\frac{d\theta}{2\pi}.$$

The right-hand side is $O(1)$ by basic properties of the Gaussian distribution. (Notice that $p(z)$, for any fixed z is a Gaussian random variable.) This proves (26). Using the bound $|\mathbf{E}p(z)| = O(t)$ for all $z \in B(x,2\gamma)$, one concludes that with probability at least $1-O(\gamma^{16/10}\delta_n^{2c_1})$,

$$\max_{z\in\partial B(x,2\gamma)}|p(z)| \leq Kt, \tag{27}$$

for some constant $K > 0$.

Now, we address $g(z)$; since g is a linear function with real coefficients, we have

$$\min_{z\in\partial B(x,2\gamma)}|g(z)| = \min\{|g(x-2\gamma)|,|g(x+2\gamma)|\},$$

which reduces the task to obtaining lower bounds for the two end points only.

Note that $g(x+2\gamma)$ is normally distributed with standard deviation

$$\sqrt{\mathrm{Var}\left(g(x+2\gamma)\right)} = \sqrt{\sum_{j=1}^n\left(\phi_j(x)+2\gamma\phi_j'(x)\right)^2} \geq \sqrt{\sum_{j=1}^n\phi_j^2(x)} - 2\gamma\sqrt{\sum_{j=1}^n\phi_j'^2(x)} \geq 1/2\sqrt{\sum_{j=1}^n\phi_j^2(x)}$$

wherein the last two inequalities, we used the triangle inequality and then Condition C2(5). Note that by the definition of t,

$$\sqrt{\sum_{j=1}^n\phi_j^2(x)} = \sqrt{\mathrm{Var}\left(\tilde{F}_n(x)\right)} = t\delta_n^{-2c_2+c_1}.$$

Since $g(x+2\gamma)$, as a random variable, is a real Gaussian with density bounded by $\frac{1}{2\sqrt{\mathrm{Var}\,g(x+2\gamma)}} \leq \frac{\delta_n^{2c_2-c_1}}{t}$, we have for any constant $K > 0$,

$$\mathbf{P}\left(|g(x+2\gamma)| \leq Kt\right) = O\left(\delta_n^{2c_2-c_1}\right) = O\left(\delta_n^{2c_1}\gamma^{3/2}\right).$$

In the last inequality we used the fact that c_2 is set to be much larger than c_1; see the paragraph following (15).

We can prove a similar statement for $g(x-2\gamma)$. Thus we can conclude that for any constant $K>0$,

$$\mathbf{P}\left(\min_{z\in\partial B(x,2\gamma)}|g(z)|\le Kt\right)=O\big(\delta_n^{2c_1}\gamma^{3/2}\big). \tag{28}$$

Combining (28) and (27), we conclude the proof of Lemma 8.5. □

9. Proof of Theorem 3.2. In this section, we prove Theorem 3.2 by applying Theorem 2.6. By dividing the coefficients c_i and d_i by their maximum modulus, it suffices to assume that $\max_{0\le j\le n}\{|c_j|,|d_j|\}=1$. For the sake of simplicity, we assume all random variables have mean 0; the more general setting in Condition C4 can be dealt with via a routine modification.

Our crucial new ingredient is the following lemma, which is a generalization of a classical result of Turán [59].

LEMMA 9.1. [40, Chapter I] *For $i=\sqrt{-1}$, let*

$$p(t)=\sum_{k=0}^{h}a_ke^{i\lambda_kt},\quad a_k\in\mathbb{C},\ \lambda_0<\lambda_1<\cdots<\lambda_h\in\mathbb{R}.$$

Then for any interval $J\subset\mathbb{R}$ and any measurable subset $E\subset J$ of positive measure, we have

$$\max_{t\in J}|p(t)|\le\left(\frac{C|J|}{|E|}\right)^h\sup_{t\in E}|p(t)|$$

where C is an absolute constant.

We shall apply Theorem 2.6 to the function $F_{2n+1}(z):=P_n(10^4Cz/n)$ and the number of summands is $2n+1$ in place of n (so we only care about F_k where k is odd). The corresponding parameters are $\delta_{2n+1}:=1/n$, and $D_{2n+1}:=\{z:|\operatorname{Im}(z)|\le 1/10^4\}$. The functions ϕ_i in (1) are

$$\phi_1(z)=c_0,\phi_2(z)=c_1\cos(z),\dots,\phi_{n+1}(z)=c_n\cos(nz),$$
$$\phi_{n+2}(z)=d_1\sin(z),\dots,\phi_{2n+1}(z)=d_n\sin(nz)$$

and the random variables $\xi_1,\dots,\xi_{2n+1}$ in (1) will be $\xi_0,\dots,\xi_n,\eta_1,\dots,\eta_n$, respectively. The constant 10^4 is chosen rather arbitrarily, any sufficiently large constant would work.

To deduce Theorem 3.2 from Theorem 2.6, for this model, we set $\alpha_1=1/2$ and C_1 to be any constant larger than 1. We only need to show that for any positive constants A,c_1, there exists a constant C for which Condition C2 holds with parameters (k,C_1,α_1,A,c_1,C).

For Condition C2(1), notice that the periodic function P_n has at most $2n$ complex zeros in the region $[a, a+2\pi) \times \mathbb{R} \subset \mathbb{C}$ for any $a \in \mathbb{R}$. Indeed, let $w = e^{iz}$ then

$$w^n P(z) = \frac{1}{2}\left(\sum_{k=0}^{n} \xi_k \left(w^{n+k} + w^{n-k}\right) - i\sum_{k=1}^{n} \eta_k \left(w^{n+k} - w^{n-k}\right)\right)$$

which is a polynomial of degree $2n$ in w and has at most $2n$ zeros. For each w there is only one z in the above region that corresponds to w. Thus this condition holds trivially for any constant $C_1 > 1$, as the left-hand side of Condition C2(1) becomes zero.

Now we address (the critical) Condition C2(2). We will prove the following stronger statement that for every positive constants c_1, A, there exists a constant C' such that the following holds. For every complex number z_0, there exists a real number x such that $|x - z_0| \le |\text{Im}(z_0)| + \frac{1}{n}$ and

$$\mathbf{P}\big(|P(x)| \le \exp(-n^{c_1})\big) \le C' n^{-A}.$$

Let $x_0 = \text{Re}(z_0)$ and $I = [x_0 - \frac{1}{n}, x_0 + \frac{1}{n}]$. By conditioning on the random variables η_i and replacing A by $2A$, it suffices to show that there exists $x \in I$ for which

$$\sup_{Z \in \mathbb{R}} \mathbf{P}\left(\left|\sum_{j=0}^{n} c_j \xi_j \cos(jx) - Z\right| \le e^{-n^{c_1}}\right) \le C' n^{-A/2}. \tag{29}$$

Now let us recall the definition of $\mathcal{I}_0$ in Condition C3. We would like to point out that in this part of the proof, we only use the fact that the size of $\mathcal{I}_0$ is of order $\Theta(n)$.

We shall prove a more general version which will be useful for all of the remaining models in this manuscript.

LEMMA 9.2. *Let $\mathcal{E}$ be an index set of size $N \in \mathbb{N}$, and let $(\xi_j)_{j \in \mathcal{E}}$ be independent random variables satisfying the moment Condition C1(i). Let $(e_j)_{j \in \mathcal{E}}$ be deterministic (real or complex) coefficients with $|e_j| \ge \bar{e}$ for all j and for some number $\bar{e} \in \mathbb{R}_+$. Then for any $A \ge 1$, any interval $I \subset \mathbb{R}$ of length at least N^{-A}, there exists an $x \in I$ such that*

$$\sup_{Z \in \mathbb{R}} \mathbf{P}\left(\left|\sum_{j \in \mathcal{E}} e_j \xi_j \cos(jx) - Z\right| \le \bar{e} N^{-16A^2}\right) = O_A\big(N^{-A/2}\big)$$

where the implicit constant depends only on A and the constants in Condition C1(i).

Assuming this Lemma, we condition on the random variables $(\xi_j)_{j\notin\mathcal{I}_0}$ and apply the Lemma with $\mathcal{E} := \mathcal{I}_0, e_j := c_j, N := |\mathcal{I}_0| = \Theta(n)$ to obtain (29) directly with $\bar{e} = \Theta(1)$.

Proof of Lemma 9.2. We will prove Lemma 9.2 in three steps. In the first (and most important) step, we handle the case where ξ_i are iid Rademacher. In the second step, we handle the case where the ξ_i have symmetric distributions. In the final step, we address the most general setting.

Step 1. ξ_i are iid Rademacher (that is, $\mathbf{P}(\xi_i = 1) = \mathbf{P}(\xi_i = -1) = 1/2$). The key ingredient in this step is the following inequality, which is a variant of a result of Halász [25]; see also [57, Cor. 7.16], [43, Cor. 6.3] for relevant estimates. Before stating the result, we recall a definition of multi-sets: a multi-set is a collection of unordered elements in which each element can appear more than once.

LEMMA 9.3. *Let $\varepsilon_1, \dots, \varepsilon_n$ be independent Rademacher random variables. Let $a_1, \dots, a_n$ be real numbers and l be a fixed integer. Assume that there is a constant $a > 0$ such that for any two different multi-sets $\{i_1, \dots, i_{l'}\}$ and $\{j_1, \dots, j_{l''}\}$ where $l' + l'' \le 2l$, $|a_{i_1} + \cdots + a_{i_{l'}} - a_{j_1} - \cdots - a_{j_{l''}}| \ge a$. Then*

$$\sup_{Z\in\mathbb{R}} \mathbf{P}\left(\left|\sum_{j=1}^{n} a_j\varepsilon_j - Z\right| \le an^{-l}\right) = O_l\big(n^{-l}\big).$$

For the sake of completeness, we present a short proof of this lemma in Appendix 15.4.

There exists a subset $\mathcal{E}' \subset \mathcal{E}$ of size at least half the size of $\mathcal{E}$ such that either for all $i \in \mathcal{E}'$, $|\mathrm{Re}(e_i)| \ge \bar{e}/2$ or for all $i \in \mathcal{E}'$, $|\mathrm{Im}(e_i)| \ge \bar{e}/2$. Since

$$\mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\cos(jx) - Z\right| \le \bar{e}N^{-16A^2}\right) \le \mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} \mathrm{Re}(e_j)\xi_j\cos(jx) - \mathrm{Re}(Z)\right| \le \bar{e}N^{-16A^2}\right)$$

and

$$\mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\cos(jx) - Z\right| \le \bar{e}N^{-16A^2}\right) \le \mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} \mathrm{Im}(e_j)\xi_j\cos(jx) - \mathrm{Im}(Z)\right| \le \bar{e}N^{-16A^2}\right),$$

we can, by conditioning on the $(\xi_j)_{j\notin\mathcal{E}'}$ and replacing $\mathcal{E}$ by $\mathcal{E}'$, assume that the e_i are real and Z is real. This allows us to apply Lemma 9.3.

In order to apply Lemma 9.3, we first show that there exists an $x\in I$ such that for every 2 distinct multi-sets $\{i_1,\dots,i_{A'}\}$ and $\{j_1,\dots,j_{A''}\}$ in $\mathcal{E}$ with $A'+A''\le 2A$, we have

$$\left|\sum_{t=1}^{A'} e_{i_t}\cos(i_t x)-\sum_{t=1}^{A''} e_{j_t}\cos(j_t x)\right| > \bar{e}N^{-16A^2}N^A. \tag{30}$$

Let us fix such two multi-sets and let

$$h(x):=\sum_{t=1}^{A'} e_{i_t}\cos(i_t x)-\sum_{t=1}^{A''} e_{j_t}\cos(j_t x).$$

Let $E:=\{x\in I:|h(x)|\le \bar{e}N^{-16A^2}N^A\}$. Since h can be written in terms of exponential polynomials with $4A$ frequencies, we can apply Lemma 9.1 to obtain

$$\max_{[0,2\pi]}|h|\le\left(\frac{C'}{|E|}\right)^{4A}\sup_E|h|. \tag{31}$$

By the definition of E, the right-hand side is bounded from above by

$$\left(\frac{C'}{|E|}\right)^{4A}\bar{e}N^{-16A^2}N^A.$$

To bound the left-hand side from below, observe from orthogonality of the functions $\cos kx$ that

$$2\pi\max_{[0,2\pi]}|h|^2\ge\int_0^{2\pi}|h|^2dx\ge\pi\bar{e}^2, \tag{32}$$

as all $|e_i|$ with $i\in\mathcal{E}$ is at least $\bar{e}$.

Therefore, from (31), we get $|E|=O_A(N^{-4A+1/4})$. Since there are only $O(N^{2A})$ choices for the sets A' and A'', we conclude that every x in I, except for a set of Lebesgue measure at most $O_A(N^{-2A+1/4})=o_A(|I|)$, satisfies (30).

To conclude the proof, we use (30) with Lemma 9.3. By setting $a:=\bar{e}N^{-16A^2}N^A$ and $l:=A$, Lemma 9.3 gives

$$\sup_{Z\in\mathbb{C}}\mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\cos(jx)-Z\right|\le\bar{e}N^{-16A^2}\right)=O_A\big(N^{-A}\big). \tag{33}$$

This proves Lemma 9.2 for the Rademacher case.

Step 2. In this step, we consider the case where random variables ξ_j have symmetric distributions. In this case, $(\xi_j)_j$ and $(\xi_j\varepsilon_j)_j$ have the same distribution where ε_j are independent Rademacher random variables that are independent of the ξ_j. Thus, the claimed statement is equivalent to

$$\sup_{Z\in\mathbb{R}} \mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\varepsilon_j\cos(jx) - Z\right| \le \bar{e}N^{-16A^2}\right) = O_A\big(N^{-A/2}\big) \tag{34}$$

for some $x \in I$.

The natural way to prove this is to use the (standard) conditioning argument, one fixes all ξ_j and uses the Rademacher variables as the only random source, going back to Step 1. However, the situation here is more delicate, as x may not be the same in each evaluation of ξ_j. We handle this extra complication by proving the stronger statement that

$$\fint_I \sup_{Z\in\mathbb{R}} \mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\varepsilon_j\cos(jx) - Z\right| \le \bar{e}N^{-16A^2}\right) dx = O_A\big(N^{-A/2}\big) \tag{35}$$

where $\fint_I f\,dx := \frac{1}{|I|}\int_I f\,dx$.

The left-hand side is at most

$$\fint_I \mathbf{E}_{(\xi_j)} \sup_{Z\in\mathbb{R}} \mathbf{P}_{(\varepsilon_j)}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\varepsilon_j\cos(jx) - Z\right| \le \bar{e}N^{-16A^2}\right) dx.$$

By Fubini's theorem, it suffices to show that

$$\mathbf{E}_{(\xi_j)}\fint_I \sup_{Z\in\mathbb{R}} \mathbf{P}_{(\varepsilon_j)}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\varepsilon_j\cos(jx) - Z\right| \le \bar{e}N^{-16A^2}\right) dx = O_A\big(N^{-A/2}\big). \tag{36}$$

We first show that with high probability, there are $\Theta(N)$ indices $j \in \mathcal{E}$ such that $|\xi_j| = \Theta(1)$, which is needed to guarantee (32). Assume, for a moment, that $\mathbf{P}(|\xi_j| < d) \ge 1-d$ for some small positive constant d. Since the random variables ξ_j are symmetric, they have mean 0. Using the boundedness of the $(2+\varepsilon)$ central moment of ξ_j (Condition C1), and the fact that ξ_j has variance 1, we have

$$\begin{aligned}\mathbf{E}|\xi_j|^2 = 1 = \mathbf{E}|\xi_j|^2\mathbf{1}_{|\xi_j|<d} + \mathbf{E}|\xi_j|^2\mathbf{1}_{|\xi_j|\ge d} &\le d^2 + d^{\varepsilon/(2+\varepsilon)}\big(\mathbf{E}|\xi_j|^{2+\varepsilon}\big)^{2/(2+\varepsilon)}\\ &\le d^2 + d^{\varepsilon/(2+\varepsilon)}\tau^{2/(2+\varepsilon)}.\end{aligned}$$

Thus, if d is small enough (depending on τ and ε), we have a contradiction. Hence, there is a constant $d > 0$ such that $\mathbf{P}(|\xi_j| < d) \le 1-d$. Now, by Chernoff's inequality, with probability at least $1 - e^{-\Theta(N)}$, there are at least $\Theta(N)$ indices $j \in \mathcal{E}$ for which $|\xi_j| \ge d$. On the event that this happens, we condition on the ε_j where $|\xi_j| < d$ and use Step 1 to conclude that outside a subset of I of measure at most $O_A(N^{-2A+1/4})$, we have

$$\sup_{Z\in\mathbb{C}} \mathbf{P}_{(\varepsilon_j)}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\varepsilon_j\cos(jx)-Z\right|\le \bar{e}N^{-16A^2}\right)=O_A\big(N^{-A}\big).$$

Therefore, the left-hand side of (36) is at most

$$\begin{aligned} e^{-\Theta(N)}\fint_I 1dx+O_A\big(N^{-2A+1/4}/|I|\big)+O_A\left(\fint_I N^{-A}dx\right) &= O_A\big(N^{-A+1/4}\big)\\ &= O_A\big(N^{-A/2}\big), \end{aligned}$$

completing the proof for this case.

Step 3. Finally, we address the general case. Let ξ'_j be independent copies of ξ_j, $j\in\mathcal{E}$. Then the variables $\xi''_j:=\xi_j-\xi'_j$ are symmetric and have uniformly bounded $(2+\varepsilon)$-moments. By Step 2, we have

$$\begin{aligned} &\left[\mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi_j\cos(jx)-Z\right|\le \bar{e}N^{-16A^2}\right)\right]^2\\ &\quad\le \mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j\xi''_j\cos(jx)\right|\le 2\bar{e}N^{-16A^2}\right)\le O_A\big(n^{-A}\big) \end{aligned}$$

wherein the last inequality, we decompose the disk $B(0,2\bar{e}N^{-16A^2})$ into $O(1)$ disks of radius $\bar{e}N^{-16A^2}$ (not necessarily centered at 0) before applying Step 2. Taking square root of both sides, we obtain Lemma 9.2. □

The remaining conditions are easy to check. Condition C2(3) follows from the following lemma.

LEMMA 9.4. *For any positive constants* A, c_1 *and* C, *we have, with probability at least* $1-O(n^{-A})$, $\log M\le n^{c_1}$, *where* $M:=\max\{|P(z)|:|\mathrm{Im}(z)|\le C/n\}$.

Proof. For every $1\le j\le n$, we have $|e^{ijz}|=e^{-j\,\mathrm{Im}(z)}\le e^C$. And so,

$$\max_{|\mathrm{Im}(z)|\le C/n,1\le j\le n}\big\{|\cos jz|,|\sin jz|\big\}\le e^C. \tag{37}$$

Let B be the event on which $|\xi_j|\le n^{A/2+1}$ for all $0\le j\le n$. Notice that on the complement B^c of B, $\log M=o(n^{c_1})$ for any constant $c_1>0$. By Chebyshev's inequality (exploiting the fact that $\mathbf{E}|\xi_i|^2=1$) and the union bound, we have

$$\mathbf{P}(B^c)\le \frac{n}{n^{A+2}}=o\big(n^{-A}\big),$$

completing the proof. □

Finally Condition C2(4) follows from the following lemma:

LEMMA 9.5. *For any constant C, there exists a constant $C' > 0$ such that for every z with $|\mathrm{Im}(z)| \le C/n$,*

$$\frac{|c_j||\cos(jz)|}{\sqrt{S}} \le C' n^{-1/2}, \quad \textit{for all } 0 \le j \le n, \tag{38}$$

and

$$\frac{|d_j||\sin(jz)|}{\sqrt{S}} \le C' n^{-1/2}, \quad \textit{for all } 0 \le j \le n, \tag{39}$$

where $S := \sum_{j=0}^{n} |c_j|^2 |\cos jz|^2 + \sum_{j=1}^{n} |d_j|^2 |\sin jz|^2$.

Proof. Write $z =: a + ib$. Without loss of generality, assume that $b \ge 0$. By (37), $|\cos(jz)| \le C$ and $|\sin(jz)| \le C$ for all $0 \le j \le n$, so it suffices to show $S := \Omega(n)$. To achieve this bound on S, it suffices to show that $\mathcal{I}_0$ contains a subset J of size $\Theta(n)$ such that

$$|\cos(jz)| \ge c^* \quad \text{for all } j \in J \text{, for some positive constant } c^*. \tag{40}$$

Since $b \ge 0$ and $j \ge 0$, we have

$$2|\cos(jz)| = e^{jb} |e^{-2jb+2ija} + 1| \ge |w^j + 1|$$

where $w := e^{-2b+2ia}$. By Condition C3 and Lemma 3.1, we can find a subset J of $\mathcal{I}_0$ of size $\Theta(n)$ such that

$$\min_{k \in \mathbb{Z}} \{ |2aj - (2k+1)\pi| \} \ge c$$

for some constant $c > 0$ and all $j \in J$. We can assume, without loss of generality, that $c \le 1/10$ and this guarantees $|\cos(2aj) + 1| \ge c^2/4$.

Consider $j \in J$, if $1 - e^{-2jb} \ge c^2/10$ then by the triangle inequality,

$$|w^j + 1| = |e^{-2jb} e^{2iaj} + 1| \ge 1 - |e^{-2jb} e^{2iaj}| = |e^{-2jb} - 1| \ge c^2/10.$$

In the opposite case, $e^{-2jb} \ge 1 - c^2/10 > .99$. Keeping in mind that $c \le 1/10$, we have

$$|w^j + 1| \ge e^{-2jb} |e^{2iaj} + 1| - |e^{-2jb} - 1| \ge .99c^2/4 - c^2/10 \ge c^2/10. \tag{41}$$

Thus, we achieved (40) with $c^* = c^2/10$. □

Finally, using Conditions C3, C4 and (40), it is a routine to prove that the repulsion Condition C2(5) holds. That completes the proof.

10. Proof of Theorem 3.5 and Corollary 3.6. As before, by rescaling the coefficients, we can assume that $\max_{0\le j\le n}\{|c_j|,|d_j|\}=1$. Before going to the proofs, let us state a version of the Kac-Rice formula for Gaussian processes. Note that a Gaussian process $P(t), t\in(a_0,b_0)$ is a random variable $P:\Omega\times(a_0,b_0)\to\mathbb{R}$ with Ω being a probability space such that for each $\omega\in\Omega$, $P(\omega,\cdot)$ is a continuous function on (a_0,b_0) and for each $k\in\mathbb{N}$, $t_1,\dots,t_k\in(a_0,b_0)$, $(P(\cdot,t_1),\dots,P(\cdot,t_k))$ is a Gaussian random vector.

PROPOSITION 10.1. [17, Theorem 2.5] *Let $P(t)$, $t\in(a_0,b_0)$ be a real, differentiable Gaussian process. Let*

$$\mathcal{P}(t)=\mathrm{Var}\big(P(t)\big),\ \mathcal{Q}(t)=\mathrm{Var}\big(P'(t)\big),\ \mathcal{R}(t)=\mathrm{Cov}\big(P(t),P'(t)\big),$$

$$\rho(t)=\frac{\mathcal{R}(t)}{\sqrt{\mathcal{P}(t)\mathcal{Q}(t)}},\ m(t)=\mathbf{E}P(t),\ \text{and}\ \eta(t)=\frac{m'(t)-\rho(t)m(t)\sqrt{\mathcal{Q}(t)/\mathcal{P}(t)}}{\sqrt{\mathcal{Q}(t)(1-\rho^2(t))}}.$$

Assume that $m'(t)$ is continuous and the joint normal distribution for $P(t)$ and $P'(t)$ has non-singular covariance matrix for each t, then for any interval $[a,b]\subset(a_0,b_0)$, we have

$$\mathbf{E}N_P(a,b)$$
$$=\int_a^b\sqrt{\frac{\mathcal{Q}(t)\big(1-\rho^2(t)\big)}{\mathcal{P}}}\phi\left(\frac{m(t)}{\sqrt{\mathcal{P}(t)}}\right)\big(2\phi\big(\eta(t)\big)+\eta(t)\big(2\Phi\big(\eta(t)\big)-1\big)\big)dt$$

where $\phi(t)$ and $\Phi(t)$ are the standard normal density and distribution functions, respectively.

Proof of Theorem 3.5. By triangle inequality, we can assume that $\tilde{\xi}_j$ and $\tilde{\eta}_j$ are Gaussian random variables. Let c be the constant in Theorem 3.2 with $\alpha_1=1/2, k=1, l=0$. As in Remark 2.7, we can set $c=\frac{\varepsilon}{2\cdot 10^9}$. Let $\alpha=c/7$. It suffices to show that for every interval (a_n,b_n) of size at most $1/n$, we have

$$\big|\mathbf{E}N_{P_n}(a_n,b_n)-\mathbf{E}N_{\tilde{P}_n}(a_n,b_n)\big|=O\big(n^{-\alpha/2}\big). \tag{42}$$

If $b_n-a_n\ge 1/n$, we simply divide the interval (a,b) into $\lfloor(b_n-a_n)n\rfloor+1$ intervals of size at most $1/n$ each and then apply (42) to each interval and then sum up the bounds.

Let $\ell:=(b_n-a_n)/2$. Let G be a smooth function on $\mathbb{R}$ with support in $\left[\frac{a_n+b_n}{2}-\ell-n^{-1-\alpha},\frac{a_n+b_n}{2}+\ell+n^{-1-\alpha}\right]$ such that $0\le G\le 1$, $G=1$ on $\left[\frac{a_n+b_n}{2}-\ell,\frac{a_n+b_n}{2}+\ell\right]$, and $\left\|G^{(a)}\right\|_\infty\le Cn^{6\alpha+a}$ for all $0\le a\le 6$.

By the definition of G, we have

$$\mathbf{E}N_{P_n}(a_n,b_n)\le\mathbf{E}\sum G(\zeta_i)\le\mathbf{E}N_{P_n}\big(a_n-n^{-1-\alpha},b_n+n^{-1-\alpha}\big)$$

where ζ_i are the real roots of P_n. Similarly,

$$\mathbf{E}N_{\tilde{P}_n}(a_n,b_n) \le \mathbf{E}\sum G(\tilde{\zeta}_i) \le \mathbf{E}N_{\tilde{P}_n}\big(a_n - n^{-1-\alpha}, b_n + n^{-1-\alpha}\big).$$

Applying Theorem 3.2 (with $k=1, l=0$) to the function $G/n^{6\alpha}$, we get

$$\mathbf{E}\sum G(\zeta_i) = \mathbf{E}\sum G(\tilde{\zeta}_i) + O\big(n^{-c+6\alpha}\big) = \mathbf{E}\sum G(\tilde{\zeta}_i) + O\big(n^{-\alpha}\big).$$

Since $\alpha = c/7$, we obtain

$$\begin{aligned}\mathbf{E}N_{P_n}(a_n,b_n) &\le \mathbf{E}N_{\tilde{P}_n}(a_n - n^{-1-\alpha}, b_n + n^{-1-\alpha}) + O\big(n^{-\alpha}\big)\\ &\le \mathbf{E}N_{\tilde{P}_n}(a_n,b_n) + 2\mathcal{I}_{\tilde{P}_n} + O\big(n^{-\alpha}\big),\end{aligned}$$

where $\mathcal{I}_{\tilde{P}_n} := \sup_{x\in\mathbb{R}} \mathbf{E}N_{\tilde{P}_n}(x - n^{-1-\alpha}, x)$. We will show later that $\mathcal{I}_{\tilde{P}_n} = O(n^{-\alpha/2})$, which gives the upper bound $\mathbf{E}N_{P_n}(a_n,b_n) \le \mathbf{E}N_{\tilde{P}_n}(a_n,b_n) + O(n^{-\alpha/2})$.

Let us quickly address the lower bound $\mathbf{E}N_{P_n}(a_n,b_n) \ge \mathbf{E}N_{\tilde{P}_n}(a_n,b_n) - O(n^{-\alpha/2})$. If $\ell > n^{-1-\alpha}$, we can argue as for the upper bound. In the case $\ell \le n^{-1-\alpha}$, the desired bound follows from the observation that $\mathbf{E}N_{P_n}(a_n,b_n) \ge 0 \ge \mathcal{I}_{\tilde{P}_n} - O(n^{-\alpha/2}) \ge \mathbf{E}N_{\tilde{P}_n}(a_n,b_n) - O(n^{-\alpha/2})$. The upper and lower bounds together give (42).

To prove the stated bound on $\mathcal{I}_{\tilde{P}_n}$, we use Proposition 10.1, which asserts that for every $x \in \mathbb{R}$,

$$\mathbf{E}N_{\tilde{P}_n}[x - n^{-\alpha-1}, x] \le \int_{x-n^{-\alpha-1}}^{x} \sqrt{\frac{\mathcal{S}}{\mathcal{P}^2}}dt + \int_{x-n^{-\alpha-1}}^{x} \frac{|m'|\mathcal{P} + |m|\mathcal{R}}{\mathcal{P}^{3/2}}dt, \tag{43}$$

where

- $m(t) := \mathbf{E}\tilde{P}_n(t)$
- $\mathcal{P}(t) := \mathrm{Var}(\tilde{P}_n) = \sum_{k=0}^n c_k^2 \cos^2(kt) + d_k^2 \sin^2(kt)$
- $\mathcal{Q}(t) := \mathrm{Var}(\tilde{P}_n') = \sum_{k=0}^n k^2 c_k^2 \sin^2(kt) + k^2 d_k^2 \cos^2(kt)$
- $\mathcal{R}(t) := \mathrm{Cov}(\tilde{P}_n, \tilde{P}_n') = \sum_{k=0}^n k\cos(kt)\sin(kt)(-c_k^2 + d_k^2)$
- $\mathcal{S}(t) = \mathcal{P}(t)\mathcal{Q}(t) - \mathcal{R}^2(t)$.

Observe that the covariance matrix of $(P_n(t), P_n'(t))$ is non-singular if and only if $\mathcal{S}(t) \ne 0$. Since the deterministic function $\mathcal{S}(t)$ only has finitely many zeroes in $[x - n^{-\alpha-1}, x] + (-\varepsilon^*, \varepsilon^*)$ (where we add $(-\varepsilon^*, \varepsilon^*)$ only to make the interval bigger to apply Proposition 10.1, ε^* can be any positive number), we can decompose this interval into subintervals whose interiors do not contain any zero of $\mathcal{S}$, and use linearity of expectation if necessary. This way, we can assume that the joint distribution of $\tilde{P}_n$ and $\tilde{P}_n'$ is non-singular, as required in Proposition 10.1.

From (37) and (40), there is a constant $K > 0$ such that for every $t \in \mathbb{R}$,

$$\mathcal{P} \ge \frac{n}{K}, \quad \mathcal{Q} \le Kn^3, \quad \text{and} \quad \mathcal{R} \le Kn^2 \le Kn\mathcal{P}.$$

From here, we obtain (for all t) that $\frac{\mathcal{S}}{\mathcal{P}^2} \le \frac{\mathcal{Q}}{\mathcal{P}} \le Kn^2$.

Moreover, from Condition C4, we have $|m(t)| \le Kn^{\tau_0}$ and $|m'(t)| \le Kn^{1/2+\tau_0}$ (notice that $m(t) = 0$ if all atom random variables have zero mean; the upper bounds here come from the bound on the expectations). It follows that

$$\frac{|m'|\mathcal{P} + |m|\mathcal{R}}{\mathcal{P}^{3/2}} \le Kn^{1/2+\tau_0}.$$

Using the above estimates, we conclude that the integrand on the right-hand side of (43) is bounded (in absolute value) by $O(n^{1/2+\tau_0})$. Since the length of the interval in the integration is $n^{-\alpha-1}$, the integral is of order $O(n^{\tau_0-\alpha-1/2}) = O(n^{-\alpha/2})$, as $\tau_0 - 1/2 = \frac{\varepsilon}{10^{11}} \le \alpha/2$. □

Proof of Corollary 3.6. As promised in Remark 3.7, we will prove the desired statement for P_n as in (5). Applying Theorem 3.5 with

$$\tilde{P}_n(x) := u_n \sqrt{\sum_{i=0}^n c_i^2} + \sum_{j=0}^{N_0} u_j n^{1/2-\alpha} \cos(jx) + \sum_{j=1}^{N_0} v_j n^{1/2-\alpha} \sin(jx)$$
$$+ \sum_{j=0}^n c_j \tilde{\xi}_j \cos(jx) + \sum_{j=1}^n c_j \tilde{\eta}_j \sin(jx)$$

where $\tilde{\xi}_j$ and $\tilde{\eta}_j$ are iid standard Gaussian, it suffices to prove that the desired estimate holds for $\tilde{P}_n$. Applying Proposition 10.1 to $\tilde{P}_n$, we obtain

$$\mathbf{E}N_{\tilde{P}_n}(a_n, b_n)$$
$$= \int_{a_n}^{b_n} \sqrt{\frac{\sum_{i=0}^n c_i^2 i^2}{\sum_{i=0}^n c_i^2}} \phi\left(\frac{m(x)}{\sqrt{\sum_{i=0}^n c_i^2}}\right) \left[2\phi(q(x)) + q(x)\big(2\Phi\big(q(x)\big) - 1\big)\right] dx$$

where

$$m(x) := u_n \sqrt{\sum_{i=0}^n c_i^2} + \sum_{j=0}^{N_0} u_j n^{1/2-\alpha} \cos(jx) + \sum_{j=1}^{N_0} v_j n^{1/2-\alpha} \sin(jx)$$

and $q(x) := \frac{m'(x)}{\sqrt{\sum_{i=0}^n c_i^2 i^2}}$.

In our setting, $\sum_{i=0}^n c_i^2 = \Theta(n)$, $\sum_{i=0}^n c_i^2 i^2 = \Theta(n^3)$, and so $\frac{m(x)}{\sqrt{\sum_{i=0}^n c_i^2}} = u_n + O(n^{-\alpha})$ and $q(x) = O(n^{-1})$. Therefore, by the boundedness of the functions Φ, ϕ and ϕ', we get

$$\phi\left(\frac{m(x)}{\sqrt{\sum_{i=0}^n c_i^2}}\right) = \phi(u_n) + O\big(n^{-\alpha}\big),$$

and

$$2\phi(q(x))+q(x)\big(2\Phi\big(q(x)\big)-1\big)=2\phi(0)+O\big(n^{-1}\big).$$

It follows that

$$\begin{aligned}\mathbf{E}N_{\tilde{P}_n}(a_n,b_n)&=2\sqrt{\frac{\sum_{i=0}^n c_i^2 i^2}{\sum_{i=0}^n c_i^2}}(b_n-a_n)\phi(u_n)\phi(0)\\&\quad+O\left(n^{-\alpha}\sqrt{\frac{\sum_{i=0}^n c_i^2 i^2}{\sum_{i=0}^n c_i^2}}(b_n-a_n)\right)\\&=2\sqrt{\frac{\sum_{i=0}^n c_i^2 i^2}{\sum_{i=0}^n c_i^2}}(b_n-a_n)\phi(u_n)\phi(0)\\&\quad+O\big(n^{-\alpha}(b_n-a_n)n\big).\end{aligned}$$

Plugging in $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, we obtain

$$\mathbf{E}N_{P_n}(a_n,b_n)=\frac{b_n-a_n}{\pi}\sqrt{\frac{\sum_{j=0}^n c_j^2 j^2}{\sum_{j=0}^n c_j^2}}\exp\left(-\frac{u_n^2}{2}\right)+O\big(n^{-c}\big((b_n-a_n)n+1\big)\big)$$

where the positive constant c and the implicit constant depend only on α, N_0, K, τ_1, ε, completing the proof. □

11. Proof of Theorem 4.3 and Corollary 4.5.

Proofs of Theorem 4.3. Let us first consider the case $0<\theta_n<\frac{1}{K}$ for some sufficiently large constant $K>0$. Let $\delta_n=\theta_n+1/n$.

We apply Theorem 2.6 to the random function $F_n(z):=P_n(z\theta_n/10)$ and the domain $D_n:=\{z:1-2\theta_n\le|z\theta_n/10|\le 1-\theta_n+1/n\}$.

For this model, one can choose $\alpha_1=1/2$ and $C_1=1$. The main task is to show that for any positive constants A,c_1, there exists a constant C for which Conditions C2(1)–C2(5) hold with parameters $(k+l,C_1,\alpha_1,A,c_1,C)$. Conditions C2(4) and C2(5) can be checked by a simple algebraic manipulation, which we leave as an exercise. To verify Condition C2(3), notice that for any $M>2$, if we condition on the event Ω' on which $|\xi_i|\le M(1+\delta_n/2)^i$ for all i, then for all $z\in D_n+B(0,2)$,

$$|F_n(z)|=O(M)\sum_{i=0}^n\big(1+\delta_n/2\big)^i(1-\delta_n+2/n)^i=O\big(M\delta_n^{-1}\big).\tag{44}$$

Thus, for every $M>2$, we have

$$\mathbf{P}\big(|F_n(z)|=O\big(M\delta_n^{-1}\big)\big)=1-O\left(\sum_{i=0}^n\frac{1}{M(1+\delta_n/2)^i}\right)=1-O\left(\frac{1}{M\delta_n}\right).\tag{45}$$

Setting $M=\delta_n^{-A-1}$, we obtain Condition C2(3).

To prove Condition C2(2), we show that for any constants A and $c_1>0$, there exists a constant $B>0$ such that the following holds. For every z_0 with $1-2\theta_n\le |z_0|\le 1-\theta_n+1/n$, there exists $z=z_0e^{i\theta}$ where $\theta\in[-\delta_n/100,\delta_n/100]$ such that for every $1\le M\le n\delta_n$,

$$\mathbf{P}\big(|P_n(z)|\le e^{-\delta_n^{-c_1}}e^{-BM}\big)\le \frac{B\delta_n^A}{M^A}. \tag{46}$$

Setting $M=1$, we obtain Condition C2(2).

By writing $z_0=re^{i\theta_0}$, the bound (46) follows from a more general anti-concentration bound: there exists $\theta\in I:=[\theta_0-\delta_n/100,\theta_0+\delta_n/100]$ such that

$$\sup_{Z\in\mathbb{C}}\mathbf{P}\big(\big|P_n\big(re^{i\theta}\big)-Z\big|\le e^{-\delta_n^{-c_1}}e^{-BM}\big)\le \frac{B\delta_n^A}{M^A}.$$

Since the probability of being confined in a complex ball is bounded from above by the probability of its real part being confined in the corresponding interval on the real line, it suffices to show that

$$\sup_{Z\in\mathbb{R}}\mathbf{P}\left(\left|\sum_{j=0}^{M\delta_n^{-1}/2}\xi_j r^j\cos j\theta - Z\right|\le e^{-\delta_n^{-c_1}}e^{-BM}\right)\le \frac{B\delta_n^A}{M^A}.$$

This is, in turn, a direct application of Lemma 9.2 with $N:=M\delta_n^{-1}/2$ and $\bar{e}:=e^{-2M}\le r^j$ for all $0\le j\le M\delta_n^{-1}/2$.

Finally, to prove Condition C2(1), from (45), (46), and Jensen's inequality, we get for every $1\le M\le n\delta_n$

$$\mathbf{P}\big(N\ge \delta_n^{-c_1}+BM\big)=O\left(\frac{\delta_n^A}{M^A}\right)$$

where $N=N_{F_n}B(w,2)$, $w\in D_n$.

Let $A=k+l+2$, $c_1=1$ and $M=1,2,2^2,\dots,2^m$ where m is the largest number such that $2^m\le n\delta_n$. Combining the above inequality with the fact that $N\le n$ a.s., we get

$$\begin{aligned}\mathbf{E}N^{k+l+2}\mathbf{1}_{N\ge\delta_n^{-1}}&\le C\sum_{i=1}^m\big(\delta_n^{-1}+B2^{i+1}\big)^{k+l+2}\frac{\delta_n^A}{2^{iA}}+Cn^{k+l+2}\frac{\delta_n^A}{2^{mA}}\\&\le C\delta_n^{A-k-l-2}=O(1).\end{aligned}$$

This proves Condition C2(1) and completes the proof for $\theta_n\le 1/K$. For $\theta_n\ge 1/K$, note that Jensen's inequality implies that

$$N_{P_n}B(0,1-1/K)=O_K(1)\log\frac{\max_{w\in B(0,1-1/2K)}|P_n(w)|}{\max_{w\in B(1-1/K,1/3K)}|P_n(w)|}.$$

Thus, using the bounds (44), (45), (46) for $\theta_n = 1 - 1/K$, we get for every $1 \leq M \leq n/K$,

$$\mathbf{P}\big(N_{P_n} B(0, 1 - 1/2K) \geq BM\big) = O\left(\frac{1}{M^A}\right).$$

And so, $\mathbf{E} N_{P_n} B(0, 1 - 1/2K) = O(1)$. The same holds for $\tilde{P}_n$ and therefore the desired result follows. □

Proof of Corollary 4.5. Without loss of generality, we can assume that $\tilde{\xi}_0, \dots, \tilde{\xi}_n$ are standard Gaussian random variables. As in Remark 4.4, it suffices to restrict to the roots in the interval $[-1, 1]$. Divide this interval into $I_0 = \{x : |x| \leq 1 - 1/C\}$ and $I_1 = [-1, 1] \setminus I_0$ and denote by $N(0)$ and $N(1)$ the number of real roots of P_n in these sets, respectively. We have seen in the proof of Theorem 4.3 that $\mathbf{E} N(0) = O(1)$, and so is $\tilde{N}(0)$ which is the corresponding term for $\tilde{P}_n$.

To get $\mathbf{E} N(1) - \mathbf{E} \tilde{N}(1) = O(1)$, we decompose the interval I_1 into dyadic intervals $\pm[1 - 1/C, 1 - 1/2C), \pm[1 - 1/2C, 1 - 1/4C), \dots, \pm[1 - 2/n, 1 - 1/n)$, and finally $\pm[1 - 1/n, 1]$. In each of these intervals, say $[x, y)$, we show that $\mathbf{E} N_{P_n}[x, y) - \mathbf{E} N_{\tilde{P}_n}[x, y) = O((1 - y + 1/n)^c)$ for some positive constant c. This can be routinely done by approximating the indicator function on the interval $[x, y)$ by a smooth function and applying Theorem 4.3. We omit the details as it is similar to the proof of Theorem 3.5. □

12. Proof of Theorems 5.1 and Corollary 5.2.

Proof of Theorem 5.1. Notice that by the Borel-Cantelli lemma, with probability 1, there are only a finite number of i such that $|\xi_i| \geq 2^i$. Thus with probability 1, the radius of convergence of the series P is infinity and so P is an entire function.

A natural idea is to apply Theorem 2.5 with $n = \infty$ to the function $F_n(z) := P(z)$, with $\delta_n := |z_0|^{-1}$ and $D_n := \{z_0\}$. (We will skip the redundant subscript n in the rest of the proof.) However, since $\operatorname{Var} P(z) = e^{|z|^2}$, $|P(z)|$ is likely to be of order $\Theta(e^{|z|^2/2})$ in which case Condition C2(3) fails. The idea here is to find a proper scaling, which, at the same time, preserves the analyticity of F. We set

$$F(z) := \frac{P(z)}{e^{|z_0|^2/2} e^{(z - z_0)\bar{z}_0}}. \tag{47}$$

A routine calculation shows that $\operatorname{Var} F(z) = \Theta(1)$.

Furthermore, F is analytic and has the same roots as P. For this model, let $\alpha_1 = 1/2$ and $C_1 = 2$. The main task is to show that for any positive constants A, c_1, there exists a constant C for which Conditions C2(1)–C2(4) hold with parameters $(k, C_1, \alpha_1, A, c_1, C)$. We can, without loss of generality, assume that $|z_0|$ is sufficiently large because by Jensen's inequality, one can show that the expected number of roots of both P and $\tilde{P}$ in $B(0, K)$, for any constant K, is $O_K(1)$.

Condition C2(3) is a direct consequence of the following lemma.

LEMMA 12.1. *For any constant $A>0$, there is a constant $K>0$ such that for any $M\geq 2$,*

$$\mathbf{P}\left(\max_{z\in B(z_0,2)}|F(z)|\geq KM^A\delta^{-A-2}\right)\leq\frac{K\delta^A}{M^A}. \tag{48}$$

Proof. Let $L=|z_0|+1=\Theta(\delta^{-1})$. Let Ω' be the event that $|\xi_i|\leq M^AL^A\big(1+\frac{1}{(L+2M)^2}\big)^i$ for all $i\geq 0$. Consider its complement Ω'^c,

$$\mathbf{P}\big(\Omega'^c\big)=O\left(\sum_{i=0}^{\infty}\frac{1}{M^{2A}L^{2A}\big(1+(L+2M)^{-2}\big)^{2i}}\right)=O\left(\frac{\delta^A}{M^A}\right). \tag{49}$$

On the other hand, once Ω' holds, then for every $z\in B(z_0,2)$,

$$|P(z)|\leq\sum_{i=0}^{\infty}\frac{|\xi_i||z|^i}{\sqrt{i!}}\leq M^AL^A\sum_{i=0}^{\infty}\frac{\big(|z|+|z|^{-1}\big)^i}{\sqrt{i!}}=M^AL^AS(w)$$

where $w=|z|+|z|^{-1}$ and $S(w):=\sum_{i=0}^{\infty}\frac{w^i}{\sqrt{i!}}$. Let $x:=x(w)=\lfloor w^2-1\rfloor$. We split into the sum of $S_1:=\sum_{i=0}^{5x-1}\frac{w^i}{\sqrt{i!}}$ and $S_2:=\sum_{i=5x}^{\infty}\frac{w^i}{\sqrt{i!}}$. Since the terms $\frac{w^i}{\sqrt{i!}}$ are increasing with i running from 0 to x and then decreasing with i running from x to ∞, we have $S_1\leq 5x\frac{w^x}{\sqrt{x!}}$. Moreover,

$$|S_2|\leq\frac{w^{5x}}{\sqrt{(5x)!}}\sum_{i=0}^{\infty}\frac{w^i\sqrt{(5x)!}}{\sqrt{(i+5x)!}}\leq\frac{w^{5x}}{\sqrt{(5x)!}}S.$$

By Stirling's formula (and the fact that x is sufficiently large)

$$\frac{w^{5x}}{\sqrt{(5x)!}}\leq\sqrt{\frac{(x+2)^{5x}e^{5x}}{(5x)^{5x+1/2}}}\leq\frac{1}{2}.$$

Hence, $S_2\leq\frac{1}{2}S$, which implies

$$S\leq 2S_1\leq 10x\frac{w^x}{\sqrt{x!}}\leq 100w^2e^{w^2/2}=O\big(L^2e^{|z|^2/2}\big).$$

Thus, on Ω',

$$|P(z)|=O\big(M^AL^{A+2}e^{|z|^2/2}\big).$$

By the definition of F,

$$|F(z)|=O\left(\frac{M^AL^{A+2}e^{|z|^2/2}}{e^{|z_0|^2/2}e^{\mathrm{Re}((z-z_0)\bar{z}_0)}}\right)=O\big(M^AL^{A+2}\big)$$

which, together with (49), yield the desired claim. $\square$

Write $z_0 = re^{i\theta_0}$. To verify Condition C2(2), the idea is to apply Lemma 9.2 to the entire function

$$P\big(z_0 e^{i\theta}\big) = \sum_{j=0}^{\infty} \frac{r^j}{\sqrt{j!}} \xi_j e^{ij(\theta+\theta_0)}.$$

Note that when $|\theta| \le .01 r^{-1}$, $z_0 e^{i\theta} \in B(z_0, 1/100)$. Let $x_0 = \lfloor |z_0|^2 - 1 \rfloor$. For any $M \ge r$, we apply Lemma 9.2 to the set $\mathcal{E} = \{x_0, x_0+1, \dots, x_0+M\}$, the random variables $(\xi_j)_{j\in\mathcal{E}}$, the coefficients $e_j = \frac{r^j}{\sqrt{j!}}$ and obtain that for any positive constant $A \ge 3$, for the interval $I = [-M^{-A}, M^{-A}] \subset [-.01r^{-1}, .01r^{-1}]$, there exists $\theta \in I$ such that

$$\sup_{Z\in\mathbb{C}} \mathbf{P}\left(\left| \sum_{j\in\mathcal{E}} e_j \xi_j \cos(j\theta + j\theta_0) - Z \right| \le e_{x_0+M} M^{-16A^2} \right) = O\big(M^{-A/2}\big)$$

where we use the fact that $e_{x_0} \ge e_{x_0+1} \ge \cdots \ge e_{x_0+M}$.

This together with the assumption that $\mathrm{Re}(\xi_0), \mathrm{Im}(\xi_0), \mathrm{Re}(\xi_1), \mathrm{Im}(\xi_1), \dots$ are independent imply that

$$\sup_{Z\in\mathbb{C}} \mathbf{P}\left(\left| \sum_{j\in\mathcal{E}} e_j \xi_j \exp\big(ij(\theta+\theta_0)\big) - Z \right| \le e_{x_0+M} M^{-16A^2} \right) = O\big(M^{-A/2}\big)$$

because the distance between two complex numbers is at least the distance between their real components.

Conditioning on the random variables outside $\mathcal{E}$, we obtain some $\theta \in I$ such that with probability at least $1 - O(M^{-A/2})$,

$$\big|P\big(z_0 e^{i\theta}\big)\big| \ge e_{x_0+M} M^{-16A^2},$$

which implies

$$\begin{aligned} \big|F\big(z_0 e^{i\theta}\big)\big| &\ge \frac{e_{x_0+M} M^{-16A^2}}{\exp(r^2/2)\big|\exp\big(r^2\big(e^{i\theta}-1\big)\big)\big|} \\ &= \frac{r^{x_0+M} M^{-16A^2}}{\sqrt{(x_0+M)!}\exp(r^2/2)\big|\exp\big(r^2\big(e^{i\theta}-1\big)\big)\big|}. \end{aligned}$$

For $\theta \in I$, $|r^2(e^{i\theta}-1)| = O(r^2 M^{-A}) = O(1)$. Thus, by Stirling's formula,

$$\big|F\big(z_0 e^{i\theta}\big)\big| = \Omega\left(\frac{1}{r} \frac{r^M M^{-16A^2}}{\sqrt{(x_0+1)\cdots(x_0+M)}} \right) = \Omega\left(\frac{M^{-16A^2}}{r} \left(\frac{r^2}{r^2+M} \right)^{M/2} \right).$$

In other words, we have proved that for every constant $A \geq 3$, for every $M \geq r = |z_0|$, there exists $z \in B(z_0, 1/100)$ for which

$$\mathbf{P}\left(|F(z)| = O_A\left(\frac{M^{-16A^2}}{r}\left(\frac{r^2}{r^2+M}\right)^{M/2}\right)\right) = O_A\big(M^{-A/2}\big). \tag{50}$$

Setting $M = \lceil r \rceil$, we obtain Condition C2(2) (note that $r = \delta^{-1}$).

Combining (48) and (50) and Jensen's inequality, we get that there exists a constant K depending only on A such that for any $M \geq r$,

$$\mathbf{P}\big(N_F\big(B(z_0,1)\big) \geq M^2\big) \leq \frac{K}{M^A}.$$

Thus,

$$\begin{aligned}&\mathbf{E}N_F^{k+2}\big(B(z_0,1)\big)\mathbf{1}_{N_F(B(z_0,1))\geq r^2}\\ &\quad\leq \sum_{M=r}^{\infty} \mathbf{E}N_F^{k+2}\big(B(z_0,1)\big)\mathbf{1}_{M^2\leq N_F(B(z_0,1))\leq (M+1)^2}.\end{aligned}$$

As the right-hand side is at most $O(1)\sum_{M=r}^{\infty}\frac{K(M+1)^{2k+4}}{M^A} = O(1)$ by setting $A = 2k+6$, Condition C2(1) follows.

Finally for Condition C2(4), note that $|z|^i/\sqrt{i!}$ is maximized at $i = \lfloor |z|^2 - 1 \rfloor$. By Stirling's formula, at this i, $|z|^i/\sqrt{i!} = O\Big(\frac{\sqrt{\sum_j |z|^{2j}/j!}}{|z|^{1/2}}\Big)$. □

Proof of Corollary 5.2. As before, we simply approximate the indicator function $\mathbf{1}_B$ above and below by smooth test functions f and g whose derivatives up to order 6 are bounded by $O(r^{6a})$ for a sufficiently small constant a and $\int_{\mathbb{C}}(f - g)dm = O(r^{-a})$. Applying Theorem 5.1 to the function f, we obtain

$$\begin{aligned}\mathbf{E}N_P(B) \leq \mathbf{E}\sum_{\zeta:P(\zeta)=0} f(\zeta) &= \mathbf{E}\sum_{\tilde\zeta:\tilde P(\tilde\zeta)=0} f(\tilde\zeta) + O\big(r^{-c+6a}\big)\\ &= \mathbf{E}N_{\tilde P}(B) + O\big(r^{-a} + r^{-c+6a}\big)\end{aligned}$$

where c is the constant in Theorem 5.1. By choosing $a = c/12$, we get $\mathbf{E}N_P(B) = \mathbf{E}N_{\tilde P}(B) + O(r^{c/12})$. And similarly, applying Theorem 5.1 to the function g, we get the corresponding lower bound. This completes the proof. □

13. Proof of Theorem 6.2 and Corollary 6.4.

Proof of Theorem 6.2. We have $\operatorname{Var} P_n(z) = (|z|^2+1)^n$. As in the proof of Theorem 5.1, we will apply the framework in Section 2 to the function

$$F_n(z) = \frac{P_n(z/\sqrt{n})}{\big(|x_0|^2+1\big)^{n/2}\exp\big(\frac{n(z/\sqrt{n}-x_0)\bar{x}_0}{(|x_0|^2+1)}\big)},$$

$\delta_n = n^{-1}$ and $D_n = \{\sqrt{n}x_0\}$. We have $\text{Var}\,F(z) = \Theta(1)$. Note that the denominator is chosen so that $\text{Var}\,F(z) = \Theta(1)$, F is analytic, and $F(z) = 0$ if and only if $P(z/\sqrt{n}) = 0$. We will first show that Theorem 2.5 holds, and then we show that the conclusion of Theorem 2.6 also holds. For Theorem 2.5, it suffices to show that there exist positive constants C_1, α_1 such that for any positive constants A, c_1, there exists a constant C for which Conditions C2(1)–C2(4) hold with parameters C_1, α_1, A, c_1, C. For this model, one can choose $\alpha_1 = \varepsilon/4$ and $C_1 = 1$. Condition C2(3) follows from the following. For any constants $A, c_1 > 0$, we have

$$\mathbf{P}\left(\max_{z\in B(\sqrt{n}x_0,2)} |F(z)| \geq Ce^{n^{c_1}}\sqrt{n}\right) \leq \frac{Cn}{e^{n^{c_1}}} \tag{51}$$

for some constant C depending only on A and c_1.

Indeed, let Ω' be the event on which $|\xi_i| \leq e^{n^{c_1}}$ for all $i \geq 0$. The probability of its complement is bounded from above by

$$\mathbf{P}(\Omega'^c) \leq \frac{Cn}{e^{n^{c_1}}}.$$

On Ω', for every $z \in B(x_0, 2/\sqrt{n})$, we have

$$\begin{aligned} |P(z)| &\leq \sum_{i=0}^{n} \sqrt{\binom{n}{i}} |\xi_i||z|^i \leq e^{n^{c_1}}\sqrt{n}\sqrt{\sum_{i=0}^{n}\binom{n}{i}|z|^{2i}} \\ &= e^{n^{c_1}}\sqrt{n}\sqrt{\text{Var}\,P(z)}. \end{aligned} \tag{52}$$

Thus,

$$|F(z)| \leq Ce^{n^{c_1}}\sqrt{n}.$$

For Condition C2(2), note that the sequence $\sqrt{\binom{n}{i}}|x_0|^i$ increases from $i = 1$ to $i_0 = \lfloor 1 + \frac{(n-1)x_0^2}{1+x_0^2} \rfloor$ and then decreases. For $n^{-1/2+\varepsilon} \leq |x_0| \leq 1$, we have $\frac{n^{2\varepsilon}}{4} \leq i_0 \leq \frac{n+1}{2}$. Condition C2(2) follows by showing that for any constants $A, c_1 > 0$, there exists a constant C and an angle $\theta \in [-1/(100\sqrt{n}), 1/(100\sqrt{n})]$ such that

$$\mathbf{P}\left(|F(\sqrt{n}x_0e^{i\theta})| \leq Ce^{-n^{c_1}}\right) \leq Cn^{-A}. \tag{53}$$

We apply Lemma 9.2 to the set $\mathcal{E} = \{i_0, i_0+1, \dots, i_0+m\}$ where $m = \frac{n^{c_1/2}}{\log n}$, the random variables $(\xi_j)_{j\in\mathcal{E}}$, the coefficients $e_j = \sqrt{\binom{n}{j}}r^j$ where $r = |x_0|$, and the interval $I = [-m^{-A'}, m^{-A'}]$ where $A' = 5A/c_1$. We have

$$1 \leq \frac{e_j}{e_{j+1}} \leq \frac{\sqrt{j+1}}{r\sqrt{n-j}} \leq n^{1/2},$$

for all $j \in \mathcal{E}$, which implies

$$e_{i_0+m} \ge e_{i_0} n^{-m/2}.$$

Moreover, we have since e_{i_0} is the largest term, $\operatorname{Var} P(x_0) \le n e_{i_0}^2$, and so,

$$e_{i_0+m} \ge \frac{\sqrt{\operatorname{Var} P(x_0)}}{\sqrt{n} n^{m/2}} = \frac{\sqrt{\operatorname{Var} P(x_0)}}{\sqrt{n} e^{n^{c_1/2}}}.$$

Hence, there exists $\theta \in I$ such that for all $Z \in \mathbb{C}$,

$$\mathbf{P}\left(\left|\sum_{j\in\mathcal{E}} e_j \xi_j \cos(j\theta) - Z\right| \le \sqrt{\operatorname{Var} P(x_0)} e^{-n^{c_1/2}} m^{-16A'^2}/\sqrt{n}\right) = O\big(m^{-A'/2}\big)$$
$$= O\big(n^{-A}\big).$$

By conditioning on the random variables not in $\mathcal{E}$, we obtain

$$\mathbf{P}\left(\left|P_n\big(x_0 e^{i\theta}\big)\right| \le \sqrt{\operatorname{Var} P(x_0)} e^{-n^{c_1/2}} m^{-16A'^2}/\sqrt{n}\right) = O\big(n^{-A}\big).$$

Since $e^{-n^{c_1/2}} m^{-16A'^2}/\sqrt{n} = \Omega(e^{-n^{c_1}})$, we obtain

$$\mathbf{P}\left(\left|P_n\big(x_0 e^{i\theta}\big)\right| \le \sqrt{\operatorname{Var} P(x_0)} e^{-n^{c_1}}\right) = O\big(n^{-A}\big). \tag{54}$$

That implies (53) and therefore, Condition C2(2) follows.

Combining (51) and (53) and Jensen's inequality, we get that

$$\mathbf{P}\big(N_F\big(B\big(\sqrt{n} x_0, 1\big)\big) \ge n^{c_1}\big) \le C n^{-A}.$$

From this and the fact that $N_F(B(\sqrt{n} x_0, 1))$ is always at most n, Condition C2(1) follows.

For Condition C2(4), as we have seen above, $E_i := \sqrt{\binom{n}{i}} |x_0|^i$ is largest when $i_0 = \lfloor 1 + \frac{(n-1)x_0^2}{1+x_0^2} \rfloor \in [\frac{n^{2\varepsilon}}{4}, \frac{n+1}{2}]$. It suffices to show that the $E_{i_0} = O(n^{-\varepsilon/4}) \sqrt{\sum_i E_i^2}$ which can be deduced from showing that the consecutive terms $(E_i)_{i=i_0-n^{\varepsilon/2}}^{i_0+n^{\varepsilon/2}}$ are of the same order, i.e., $E_i/E_j = \Theta(1)$. We have for i in the above window,

$$\frac{E_{i+1}^2}{E_i^2} = \frac{|x_0|(n-i+1)}{i+1} = \Theta\left(\frac{n-i+1}{n-i_0+1} \frac{i_0+1}{i+1}\right) = \Theta\left(1 + \frac{1}{n^\varepsilon}\right).$$

Thus for all i, j in the above window,

$$\frac{E_i}{E_j} = \Theta\big(1 + \frac{1}{n^\varepsilon}\big)^{n^{\varepsilon/2}} = \Theta(1)$$

as needed. So Theorem 2.5 holds for F_n. It's left to show that the conclusion of Theorem 2.6 also holds.

Unfortunately, Condition C2(5) doesn't hold for F_n. Note that this condition is used in the proof of Theorem 2.6 only to show that (23) which says that for any $x \in [n^{-1/2+\varepsilon}, 1+n^{-1/2}]$, we have for a sufficiently small constant c,

$$\mathbf{P}\big(N_{\tilde{F}_n} B\big(\sqrt{n}x, 2n^{-c}\big) \geq 2\big) \leq Cn^{-16c/10} \tag{55}$$

where $\tilde{F}_n$ is the corresponding function with standard Gaussian coefficients.

To prove (55), we can instead use the fact that

$$\begin{aligned}
&\mathbf{P}\big(N_{\tilde{F}} B\big(\sqrt{n}x, 2n^{-c}\big) \geq 2\big) \\
&\leq \mathbf{P}\big(N_{\tilde{F}} B\big(\sqrt{n}x, 2n^{-c}\big) \cap \mathbb{C}_+ \geq 1\big) + \mathbf{P}\big(N_{\tilde{F}}\big[\sqrt{n}x - 2n^{-c}, \sqrt{n}x - 2n^{-c}\big] \geq 2\big) \\
&\leq \iint_{B(x,2n^{-c-1/2})\cap\mathbb{C}_+} \rho^{(0,1)}(z)dz + \int_{x-2n^{-c-1/2}}^{x+2n^{-c-1/2}} \int_{x-2n^{-c-1/2}}^{x+2n^{-c-1/2}} \rho^{(2,0)}(s,t)\, ds\, dt
\end{aligned}$$

where $\rho^{(0,1)}$ and $\rho^{(2,0)}$ are the $(0,1)$- and $(2,0)$-correlation functions of $\tilde{P}_n$ respectively. By [58, Proposition 13.3], these functions are bounded for all $z \in B(x, 2n^{-c-1/2}) \cap \mathbb{C}_+$ and $s,t \in [x - 2n^{-c-1/2}, x + 2n^{-c-1/2}]$ as follows

$$\rho^{(0,1)}(x,y) = O\big(n^{3/2}\big)(x-y) = O\big(n^{1-c}\big)$$

and

$$\rho^{(2,0)}(z) = O(n).$$

Thus,

$$\mathbf{P}\big(N_{\tilde{F}} B\big(\sqrt{n}x, 2n^{-c}\big) \geq 2\big) = O\big(n^{-2c}\big)$$

giving the desired estimate. □

Proof of Corollary 6.4. As mentioned in remark 6.3, it suffices to show that

$$\mathbf{E}N_{P_n}[0,1] = \frac{1}{4}\sqrt{n} + O\big(n^{1/2-c}\big).$$

We partition the interval $[0,1]$ into 2 intervals $I_1 := [0, n^{-1/2+\varepsilon}]$ and $I_2 := [n^{-1/2+\varepsilon}, 1]$. On the interval I_2 where Theorem 6.2 applies, we further partition it into equal intervals J_i of length $n^{-1/2}$. On each of these small intervals J_i, we routinely approximate its indicator function above and below by smooth test functions and apply Theorem 6.2 to these functions to obtain

$$\mathbf{E}N_{P_n}(J_i) - \mathbf{E}N_{\tilde{P}_n}(J_i) = O\big(n^{-c}\big).$$

Thus,

$$\mathbf{E}N_{P_n}(I_2) - \mathbf{E}N_{\tilde{P}_n}(I_2) = O\big(n^{1/2-c}\big).$$

It remains to show that the interval I_1 is insignificant. Note that $N_{P_n}(I_1) \le N_{P_n}B(x,3x)$ where $x = n^{-1/2+\varepsilon}$. By Jensen's inequality,

$$N_{P_n}B(x,3x) \le C\log\frac{M}{|P_n(x)|}$$

where $M = \max_{|z|\le 4x}|P_n(z)|$. By (52), on the event Ω',

$$M \le e^{n^{c_1}}\sqrt{n}\sqrt{\sum_{i=0}^{n}\binom{n}{i}|4x|^{2i}} = e^{n^{\varepsilon}}\sqrt{n}(16x^2+1)^{n/2} \le \sqrt{n}e^{n^{3\varepsilon}}.$$

Thus, $\mathbf{P}(\log M \ge n^{3\varepsilon}) \le \frac{n}{e^{n^{\varepsilon}}}$. Moreover, by (54), we have $\mathbf{P}(|P_n(x)| \le e^{-n^{\varepsilon}}) \le n^{-A}$. Combining these bounds, we get

$$\mathbf{P}\big(N_{P_n}B(x,3x) \ge Cn^{3\varepsilon}\big) \le Cn^{-2}.$$

Hence,

$$\mathbf{E}N_{P_n}B(x,3x) \le Cn^{3\varepsilon} + n.n^{-2} \le (C+1)n^{3\varepsilon}.$$

This completes the proof. □

14. Proof of Theorem 7.2 and Corollary 7.3.

Proof of Theorem 7.2. The reader may notice that this proof is quite similar to the proof of Theorem 4.3. We nonetheless present it here for the reader's convenience.

Let us first consider the case $0 < \delta < \frac{1}{K}$ for some sufficiently large constant $K > 0$.

We apply Theorem 2.6 to the random function $F(z) := P(z\delta/10)$ and the domain $D := \{z : 1-2\delta \le |z\delta/10| \le 1-\delta\}$.

For this random series, we set $\alpha_1 = \min\{1/4, \gamma/2\}$ and $C_1 = 1$. The main task is to show that for any positive constants A, c_1, there exists a constant C for which Conditions C2(1)–C2(4) hold with parameters $(k+l, C_1, \alpha_1, A, c_1, C)$.

We use the following crucial property of regularly varying coefficients.

LEMMA 14.1. [18, Theorem 5, p. 423] *If $c_k^2 = \frac{k^{\gamma-1}L(k)}{\Gamma(\gamma)}$ where $L(k)$ is a slowly varying function then*

$$\lim_{a\downarrow 0}\sum_{k=0}^{\infty} c_k^2(1-at)^{2k}(2at)^{\gamma}/L\left(\frac{1}{a}\right) = 1$$

uniformly as long as t stays in a compact subset of $(0,\infty)$.

Moreover, for any positive constant $c' > 0$, there exists a constant $C > 0$ (depending on the function L) such that $\frac{1}{Ct^{c'}} \le L(t) \le Ct^{c'}$ for all $t > 0$. This simple observation can be proven using, for example, the Karamata's representation theorem [4, Proposition 1.3.8, p. 26].

To verify Condition C2(4), we use Lemma 14.1 to get for every $w \in B(0, 1-\delta/2)$,

$$\sum_{k=0}^{\infty} c_k^2 |w|^{2k} = \Omega\big(\delta^{-\gamma} L(\delta^{-1})\big) = \Omega\big(\delta^{-\gamma+c'}\big)$$

while

$$c_k^2 |w|^{2k} \le Ck^{\gamma-1+c'}(1-\delta)^{2k} = O\big(\delta^{-\gamma+1-2c'}+1\big).$$

Letting c' sufficiently small, we obtain Condition C2(4).

Condition C2(5) follows immediately from Lemma 14.1.

To verify Condition C2(3), notice that for any $M > 2$, if we condition on the event Ω' on which $|\xi_i| \le M(1+\delta/2)^i$ for all i, then for all $z \in D + B(0,3)$, by Lemma 14.1,

$$|F(z)| = O(M)\sum_{i=0}^{\infty}\big(1+|c_i|^2\big)(1+\delta/2)^i(1-\delta)^i = O\big(M\delta^{-\gamma-1}\big). \tag{56}$$

Thus, for every $M > 2$, we have

$$\mathbf{P}\big(|F(z)| = O\big(M\delta^{-\gamma-1}\big)\big) = 1 - O\left(\sum_{i=0}^{n}\frac{1}{M(1+\delta/2)^i}\right) = 1 - O\left(\frac{1}{M\delta}\right). \tag{57}$$

Setting $M = \delta^{-A-1}$, we obtain Condition C2(3).

To prove Condition C2(2), we show that for any constants A and $c_1 > 0$, there exists a constant $B > 0$ such that the following holds. For every z_0 with $1-2\delta \le |z_0| \le 1-\delta$, there exists $z = z_0 e^{i\theta}$ where $\theta \in [-\delta, \delta]$ such that for every $M \ge 1$,

$$\mathbf{P}\big(|P(z)| \le e^{-\delta^{-c_1}} e^{-BM}\big) \le \frac{B\delta^A}{M^A}. \tag{58}$$

Setting $M = 1$, we obtain Condition C2(2).

By writing $z_0 = re^{i\theta_0}$, the bound (46) follows from a more general anti-concentration bound: there exists $\theta \in I := [\theta_0 - \delta, \theta_0 + \delta]$ such that

$$\sup_{Z\in\mathbb{C}} \mathbf{P}\big(|P\big(re^{i\theta}\big) - Z| \le e^{-\delta^{-c_1}} e^{-BM}\big) \le \frac{B\delta^A}{M^A}.$$

Since the probability of being confined in a complex ball is bounded from above by the probability of its real part being confined in the corresponding interval on the real line, it suffices to show that

$$\sup_{Z\in\mathbb{R}} \mathbf{P}\left(\left|\sum_{j=0}^{M\delta^{-1}/2} c_j\xi_j r^j \cos j\theta - Z\right| \le e^{-\delta^{-c_1}} e^{-BM}\right) \le \frac{B\delta^A}{M^A}.$$

This is a direct application of Lemma 9.2.

Finally, to prove Condition C2(1), from (45), (46), and Jensen's inequality, we get for every $1 \le M \le n\delta$

$$\mathbf{P}\big(N \ge \delta^{-c_1} + BM\big) = O\left(\frac{\delta^A}{M^A}\right)$$

where $N = N_F B(w,2)$, $w \in D$.

Setting $c_1 = 1$ and $M = 1, 2, 2^2, \ldots$, we get

$$\mathbf{E}N^{k+2}\mathbf{1}_{N\ge\delta^{-1}} \le C\sum_{i=1}^{\infty}\left(\delta^{-1} + B2^{i+1}\right)^{k+2}\frac{\delta^A}{2^{iA}} \le C\delta^{A-k-2}.$$

This proves Condition C2(1) and completes the proof for $\delta \le 1/K$. For $\delta \ge 1/K$, note that the Jensen's inequality implies that

$$N_P B(0, 1-1/K) = O_K(1)\log\frac{\max_{w\in B(0,1-1/2K)}|P(w)|}{\max_{w\in B(1-1/K,1/3K)}|P(w)|}.$$

Thus, using the bounds (44), (45), (46) for $\theta = 1-1/K$, and apply we get for every $1 \le M$,

$$\mathbf{P}\big(N_P B(0,1-1/2K) \ge BM\big) = O\left(\frac{C'}{M^A}\right).$$

And so, $\mathbf{E}N_P B(0,1-1/2K) = O(1)$. The same holds for $\tilde{P}$ and therefore desired result follows. □

Proof of Corollary 7.3. To prove the first part of Corollary 7.3, we decompose the interval $[0,r]$ into dyadic intervals $[0,1/2], [1-1/2, 1-1/4), \ldots$, and finally $\pm[1-\delta, r]$. In each of these interval, say $[x,y)$, we show that $\mathbf{E}N_P[x,y) - \mathbf{E}N_{\tilde{P}}[x,y) = O((1-y)^c)$ for some positive constant c. This can be routinely done by approximating the indicator function on the interval $[x,y)$ by a smooth function and apply Theorem 4.3. We omit the detail as it is similar to the proof of Theorem 3.5.

Thanks to the first part, to prove the second part of Corollary 7.3, it suffices to prove the corresponding statement for $\tilde{P}$ whose coefficients are Gaussian. We adapt a strategy in [20]. For any interval $[a,b] \subset \mathbb{R}$, by the Kac-Rice formula (Proposition

10.1), we have

$$\mathbf{E}N_{\tilde{P}}[a,b] = \frac{1}{\pi}\int_a^b \sqrt{f(x)}dx$$

where

$$f(x) = \frac{\left(\sum_{k=0}^{\infty} c_k^2 x^{2k}\right)\left(\sum_{k=0}^{\infty} c_k^2 k^2 x^{2k-2}\right) - \left(\sum_{k=0}^{\infty} c_k^2 k x^{2k-1}\right)^2}{\left(\sum_{k=0}^{\infty} c_k^2 x^{2k}\right)^2}.$$

Lemma 14.1 suggests that we make the transformation

$$f_n(t) := f(1-2^{-n}t).$$

Applying Lemma 14.1 to $a = 2^{-n}$ and $t \in [1,2]$, we obtain that uniformly on $x = 1-at \in [1-2^{1-n}, 1-2^n]$, as $n \to \infty$

$$\sum_{k=0}^{\infty} c_k^2 x^{2k} \sim 2^{-\gamma}(1-x)^{-\gamma} L(2^n), \sum_{k=0}^{\infty} c_k^2 k x^{2k-1}$$
$$\sim x^{-1} 2^{-\gamma-1}(1-x)^{-\gamma-1} L(2^n)\frac{\Gamma(\gamma+1)}{\Gamma(\gamma)}$$

and

$$\sum_{k=0}^{\infty} c_k^2 k^2 x^{2k-2} \sim x^{-2} 2^{-\gamma-2}(1-x)^{-\gamma-2} L(2^n)\frac{\Gamma(\gamma+2)}{\Gamma(\gamma)}$$

where $p_n \sim q_n$ means $\lim_{n\to\infty} \frac{p_n}{q_n} = 1$.

Since $\Gamma(\gamma+2) = (\gamma+1)\Gamma(\gamma+1) = \gamma(\gamma+1)\Gamma(\gamma)$, we obtain that uniformly on $t \in [1,2]$,

$$f_n(t) \sim \gamma(2^{-n}t)^{-2}/4.$$

We have

$$\mathbf{E}N_{\tilde{P}}[1-2^{1-n}, 1-2^{-n}] = \frac{1}{\pi}\int_1^2 2^{-n}\sqrt{f_n(t)}dt.$$

By uniform convergence, we obtain

$$\mathbf{E}N_{\tilde{P}}[1-2^{1-n}, 1-2^{-n}] \sim \frac{\sqrt{\gamma}\ln 2}{2\pi}.$$

Taking the Cesáro summation, we obtain

$$\frac{1}{n}\mathbf{E}N_{\tilde{P}}[0, 1-2^{-n}] = \frac{1}{n}\sum_{k=1}^{n}\mathbf{E}N_{\tilde{P}}[1-2^{1-k}, 1-2^{-k}] \sim \frac{\sqrt{\gamma}\ln 2}{2\pi}.$$

For each $r \in (0,1)$, sandwiching $\mathbf{E}N_{\tilde{P}}[0,r]$ between $\mathbf{E}N_{\tilde{P}}[0,1-2^{1-n}]$ and $\mathbf{E}N_{\tilde{P}}[0,1-2^{-n}]$ (i.e., $n-1=\lfloor -\log_2(1-r)\rfloor$), we get

$$\frac{1}{-\log(1-r)}\mathbf{E}N_{\tilde{P}}[0,r]\sim\frac{\sqrt{\gamma}}{2\pi}.$$

as desired. □

15. Appendix.

15.1. Proof of Lemma 8.1. This proof is taken from [58]. We will only prove the first part of the Lemma relating to Theorem 2.5 as the second part is similar. By translation, we can assume without loss of generality that $z_1=\cdots=z_k=0$. Suppose that we have (3) for G in the form (9). Let $r_0=1/100$. Then, for every function G supported in $\prod_{j=1}^k B(0,r_0)$ with $\|\nabla^a G\|_\infty\le 1$ for all $0\le a\le 2k+4$, we view it as a smooth function on the torus $(\mathbb{R}/(2.2r_0)\mathbf{Z})^{2k}$. Expanding G by Fourier series yields

$$G(w)=\sum_{b,c\in\mathbb{Z}^k} g_{b,c}e^{2\pi\sqrt{-1}(b\mathrm{Re}(w)+c\mathrm{Im}(w)/(2.2r_0))}, \tag{59}$$

for $w\in(\mathbb{R}/(2.2r_0)\mathbf{Z})^{2k}$, where

$$g_{b,c}=\frac{1}{(2.2r_0)^{2k}}\int_{B(0,r_0)^k} e^{-2\pi\sqrt{-1}(b\mathrm{Re}(w)+c\mathrm{Im}(w))/(2.2r_0)}G(w)\mathrm{d}w,$$

and the convergence is point-wise (by, for example, [22, Theorem 8.32]).

By integration by parts (or [22, Theorem 8.22e]), we have

$$|g_{b,c}|\le C(1+|b|+|c|)^{-2k-4},$$

where $C=C_k$.

Let $\eta:\mathbb{R}\to\mathbb{R}$ be a smooth function supported on $[-1.1r_0,1.1r_0]$ that equals 1 on $[-r_0,r_0]$ and $\|\eta\|_\infty\le 1$, and let

$$\psi_{b,c,i}=e^{2\pi\sqrt{-1}(b_i\,\mathrm{Re}(w_i)+c_i\,\mathrm{Im}(w_i)/(2.2r_0))}\eta\big(\mathrm{Re}(w_i)\big)\eta\big(\mathrm{Im}(w_i)\big),$$

and

$$G_{b,c}(w)=g_{b,c}\prod_{i=1}^k\psi_{b,c,i}(w_i).$$

Since G is supported on $[-r_0,r_0]^{2k}$, multiplying both sides of (59) by

$$\prod_{i=1}^k\eta\big(\mathrm{Re}(w_i)\big)\eta\big(\mathrm{Im}(w_i)\big),$$

we have

$$G(w) = \sum_{b,c\in\mathbf{Z}^k} G_{b,c}(w),$$

pointwise. We have that $\psi_{b,c,i}$ is supported on $B(0,2.2r_0)$ and $|\nabla^a G_{b,c}| \le C(1+|b|+|c|)^3|g_{b,c}|$, $\forall 0\le a\le 3$. We thus have for all $m\ge 1$

$$\begin{aligned}
&\left|\mathbf{E}\sum_{i_1,\dots,i_k} G_m\big(\zeta_{i_1},\dots,\zeta_{i_k}\big) - \mathbf{E}\sum_{i_1,\dots,i_k} G_m\big(\tilde\zeta_{i_1},\dots,\tilde\zeta_{i_k}\big)\right|\\
&\quad\le C\delta_n^c \sum_{b,c\in\mathbf{Z}^k}\big(1+|b|+|c|\big)^3|g_{b,c}|\\
&\quad\le C\delta_n^c \sum_{b,c\in\mathbf{Z}^k}\big(1+|b|+|c|\big)^{-2k-1} = C\delta_n^c\sum_{m=0}^{\infty}\sum_{b,c\in\mathbf{Z}^k,|b|+|c|=m}(1+m)^{-2k-1}\\
&\quad\le C\delta_n^c\sum_{m=0}^{\infty}(1+m)^{-2k-1}m^{2k-1}\le C\delta_n^c\sum_{m=1}^{\infty}m^{-2}\le C\delta_n^c
\end{aligned}$$

where $G_m = \sum_{|b|+|c|\le m} G_{b,c}$ supported in $B(0,2r_0)^k$ and we recall that the constant C may change from one equation to another. Using Condition C2(1) and the fact that $G_m\to G$ point-wise and $|G_m| = O(1)$, by dominated convergence theorem, we get

$$\lim_{m\to\infty}\mathbf{E}\sum_{i_1,\dots,i_k} G_m\big(\zeta_{i_1},\dots,\zeta_{i_k}\big) = \mathbf{E}\sum_{i_1,\dots,i_k} G\big(\zeta_{i_1},\dots,\zeta_{i_k}\big).$$

And hence the above inequalities hold for G in place of G_m, completing the proof.

15.2. Proof of Lemma 8.2. We follow ideas from [11]; 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 (see the beginning of Section 8), we have

$$N_{F_n}\big(B(w,1)\big)\le \log\frac{5}{2}\big(\log M - \log|F_n(w)|\big) < 2\big(\log M - \log|F_n(w)|\big)$$

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

From this and the assumption of Lemma 8.2, we conclude that

$$N_{F_n}\big(B(w,1)\big)\le 2\delta_n^{-c_2}. \tag{60}$$

By the pigeonhole principle, there exists a radius $1\ge r\ge 1/2$ for which F_n has no zeros in the annulus $B(w,r+\eta)\setminus B(w,r-\eta)$ where $\eta = .1\delta_n^{c_2}$. 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 (60), $m\le 2\delta_n^{-c_2}$. Define

$$f(z):=\frac{F_n(z)}{(z-\zeta_1)\cdots(z-\zeta_m)}.$$

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

$$|f(z)|\le|F_n(z)|\eta^{-m}\le\exp(\delta_n^{-c_2})\eta^{-m}.$$

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

$$\log|f(z)|\le\delta_n^{-c_2}+m\log\eta^{-1}\le 21\delta_n^{-2c_2}, \tag{61}$$

since

$$\delta_n^{-c_2}\le\delta_n^{-2c_2},\quad m\le 2\delta_n^{-c_2},\eta^{-1}=10\delta_n^{-c_2}\le e^{10\delta_n^{-c_2}}.$$

Because of the harmonicity of $\log|f|$, 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 $|F(w)|$ in the lemma and the fact that $|\zeta_i-w|\le 1$,

$$\log|f(w)|\ge\log|F_n(w)|\ge-\delta_n^{-c_2}. \tag{62}$$

Now, we make a critical use of Harnack's inequality [47, Chapter 11], which asserts that if a function G is harmonic on the open disk $B(w,R)$ and is non-negative continuous on its closure, for some $w\in\mathbb{C}$ and $R>0$, then for every $z\in B(w,r)$ with $r<R$,

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

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

$$21\delta_n^{-2c_2}-\log|f(z)|\le\frac{2r+\eta}{\eta}\big(21\delta_n^{-2c_2}-\log|f(w)|\big). \tag{63}$$

As $\eta=.1\delta_n^{c_2}$ and $r<1$, $\frac{2r+\eta}{\eta}\le 3\eta^{-1}=30\delta_n^{-c_2}$. By (62), the right-hand side is at most

$$30\delta_n^{-c_2}\times 22\delta_n^{-2c_2}=660\delta_n^{-3c_2}.$$

It follows that

$$\log|f(z)| \geq 21\delta_n^{-2c_2} - 660\delta_n^{-3c_2} \geq -660\delta_n^{-3c_2}.$$

Together with (61), we have

$$|\log|f(z)|| \leq 660\delta_n^{-3c_2} \quad \forall z \in B(w,r). \tag{64}$$

By the triangle inequality and the definition of f,

$$\begin{aligned} &\big\|\log|F_n(z)|\big\|_{L^2(B(w,r))} \\ &\leq \big\|\log|f(z)|\big\|_{L^2(B(w,r))} + \sum_{i=1}^{m} \big\|\log|z-\zeta_i|\big\|_{L^2(B(w,r))}. \end{aligned} \tag{65}$$

Notice that each of the m terms in the sum above is at most

$$\int_{B(0,2r-\eta)} \big|\log|z|\big|^2 dz, \quad \text{as } |\zeta_i| \leq r-\eta \text{ for all } i.$$

As $r < 1$, we can further upper bound it by $\int_{B(0,2)} |\log|z||^2 dz$, which is $O(1)$ (in fact, one can easily show $\int_{B(0,2)} |\log|z||^2 dz < 30$, with room to spare). Since $m \leq 2\delta_n^{-c_2}$, the right-hand side of (65) is at most

$$660\delta_n^{-3c_2} + 60\delta_n^{-c_2} \leq 720\delta_n^{-3c_2}.$$

Thus, we have

$$\big\|\log|F_n(z)|\big\|_{L^2(B(w,r))} \leq 720\delta_n^{-3c_2}$$

which implies the claim of the lemma as $r \geq 1/2$.

15.3. Proof of Lemma 8.3. To prove Lemma 8.3, we will follow the proofs in [11] and [58]. We first prove the following.

LEMMA 15.1. *Under the assumptions of Lemma 8.3, there exist constants $\alpha_2 > 0$ and $C' > 0$ such that for any $z_1, \dots, z_k \in D_n + B(0,1/10)$ and for any function $L : \mathbb{C}^k \to \mathbb{C}$ with continuous derivatives up to order 3 and $\|\nabla^a L\|_\infty \leq \delta_n^{-\alpha_2}$ for all $0 \leq a \leq 3$, we have*

$$\left|\mathbf{E}L\left(\frac{F_n(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{F_n(z_k)}{\sqrt{V(z_k)}}\right) - \mathbf{E}L\left(\frac{\tilde{F}_n(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{\tilde{F}_n(z_k)}{\sqrt{V(z_k)}}\right)\right| \leq C'\delta_n^{\alpha_2},$$

where $V(z_j) := \sum_{i=N_0}^{n} |\phi_i(z_j)|^2$ and N_0 is the constant in Condition C1.

Remark 15.2. Following the proof, one can set $\alpha_2 = \frac{\alpha_1 \varepsilon}{4}$.

Proof of Lemma 15.1. To prove this Lemma, we first observe that by replacing L by

$$L'(z_1,\dots,z_k) := L\left(z_1 + \frac{\mathbf{E}F_n(z_1)}{\sqrt{V(z_1)}},\dots,z_k + \frac{\mathbf{E}F_n(z_k)}{\sqrt{V(z_k)}}\right),$$

if necessary, we can assume that $\mathbf{E}\tilde{\xi}_i = 0$ for all i and $\mathbf{E}\xi_i = 0$ for all $i > N_0$. (See Condition C1.)

We use the Lindeberg swapping argument. Let

$$G_{i_0} = \sum_{i=1}^{i_0} \tilde{\xi}_i\phi_i(z) + \sum_{i=i_0+1}^{n} \xi_i\phi_i(z).$$

The purpose is to swap the random variables one by one. Under these notations, $G_0 = F_n$ and $G_n = \tilde{F}_n$. Put

$$I_{i_0} := \left|\mathbf{E}L\left(\frac{G_{i_0}(z_1)}{\sqrt{V(z_1)}},\dots,\frac{G_{i_0}(z_k)}{\sqrt{V(z_k)}}\right) - \mathbf{E}L\left(\frac{G_{i_0+1}(z_1)}{\sqrt{V(z_1)}},\dots,\frac{G_{i_0+1}(z_k)}{\sqrt{V(z_k)}}\right)\right|.$$

Then

$$I := \left|\mathbf{E}L\left(\frac{F_n(z_1)}{\sqrt{V(z_1)}},\dots,\frac{F_n(z_k)}{\sqrt{V(z_k)}}\right) - \mathbf{E}L\left(\frac{\tilde{F}_n(z_1)}{\sqrt{V(z_1)}},\dots,\frac{\tilde{F}_n(z_k)}{\sqrt{V(z_k)}}\right)\right| \le \sum_{i_0=0}^{n} I_{i_0}.$$

Fix $i_0 \in [N_0,n]$ and let $Y_j := \frac{G_{i_0}(z_j)}{\sqrt{V(z_j)}} - \frac{\xi_{i_0}\phi_{i_0}(z_j)}{\sqrt{V(z_j)}}$ for $1 \le j \le n$. Then, $\frac{G_{i_0+1}(z_j)}{\sqrt{V(z_j)}} = Y_j + \frac{\tilde{\xi}_{i_0}\phi_{i_0}(z_j)}{\sqrt{V(z_j)}}$. Condition on ξ_i for $i < i_0$ and $\tilde{\xi}_i$ for $i > i_0$. The Y_j's become constants; the only randomness left comes from $\xi_{i_0}, \tilde{\xi}_{i_0}$. Define

$$\hat{L} = \hat{L}_{i_0}(w_1,\dots,w_k) := L(Y_1 + w_1,\dots,Y_k + w_k).$$

By the definition of $\hat{L}$ and the assumption of the lemma, $\left\|\nabla^a \hat{L}\right\|_\infty \le C\delta_n^{-\alpha_2}$ for all $0 \le a \le 3$.

We are going to estimate

$$d_{i_0} := \left|\mathbf{E}_{\xi_{i_0},\tilde{\xi}_{i_0}}\hat{L}\left(\frac{\xi_{i_0}\phi_{i_0}(z_1)}{\sqrt{V(z_1)}},\dots,\frac{\xi_{i_0}\phi_{i_0}(z_k)}{\sqrt{V(z_k)}}\right)\right.$$
$$\left. - \mathbf{E}_{\xi_{i_0},\tilde{\xi}_{i_0}}\hat{L}\left(\frac{\tilde{\xi}_{i_0}\phi_{i_0}(z_1)}{\sqrt{V(z_1)}},\dots,\frac{\tilde{\xi}_{i_0}\phi_{i_0}(z_k)}{\sqrt{V(z_k)}}\right)\right|.$$

Let $a_{i,i_0} := \frac{\phi_{i_0}(z_i)}{\sqrt{V(z_i)}}$ and $a_{i_0} := (\sum_{i=1}^k |a_{i,i_0}|^2)^{1/2}$. Taylor expanding $\hat{L}$ around $(0,\dots,0)$, we obtain

$$\hat{L}\big(a_{1,i_0}\xi_{i_0},\dots,a_{k,i_0}\xi_{i_0}\big) = \hat{L}(0) + \hat{L}_1 + \mathrm{err}_1, \tag{66}$$

where

$$\hat{L}_1 = \frac{\mathrm{d}\hat{L}\big(a_{1,i_0}\xi_{i_0}t,\dots,a_{k,i_0}\xi_{i_0}t\big)}{\mathrm{d}t}\bigg|_{t=0} = \sum_{i=1}^k \frac{\partial \hat{L}(0)}{\partial \mathrm{Re}(w_i)} \mathrm{Re}\left(a_{i,i_0}\xi_{i_0}\right) + \sum_{i=1}^k \frac{\partial \hat{L}(0)}{\partial \mathrm{Im}(w_i)} \mathrm{Im}\left(a_{i,i_0}\xi_{i_0}\right).$$

(To avoid confusion, we use ∂ to denote a partial derivative of functions of multi-variables and d to denote a derivative of function of a single variable.)

From the bounds on the derivatives of $\hat{L}$, we have

$$\begin{aligned} |\mathrm{err}_1| &\le \sup_{t\in[0,1]} \left| \frac{1}{2} \frac{\mathrm{d}^2\hat{L}\big(a_{1,i_0}\xi_{i_0}t,\dots,a_{k,i_0}\xi_{i_0}t\big)}{\mathrm{d}t^2} \right| \\ &= O\left(\delta_n^{-\alpha_2}|\xi_{i_0}|^2 k \sum_{i=1}^k |a_{i,i_0}|^2\right) = O\big(\delta_n^{-\alpha_2}|\xi_{i_0}|^2 a_{i_0}^2\big). \end{aligned} \tag{67}$$

Similarly,

$$\hat{L}\big(a_{1,i_0}\xi_{i_0},\dots,a_{k,i_0}\xi_{i_0}\big) = \hat{L}(0) + \hat{L}_1 + \frac{1}{2}\hat{L}_2 + \mathrm{err}_2, \tag{68}$$

where $\hat{L}_2 = \frac{\mathrm{d}^2\hat{L}(a_{1,i_0}\xi_{i_0}t,\dots,a_{k,i_0}\xi_{i_0}t)}{\mathrm{d}t^2}\big|_{t=0}$ and

$$\begin{aligned} |\mathrm{err}_2| &\le \sup_{t\in[0,1]} \left| \frac{1}{6} \frac{\mathrm{d}^3\hat{L}\big(a_{1,i_0}\xi_{i_0}t,\dots,a_{k,i_0}\xi_{i_0}t\big)}{\mathrm{d}t^3} \right| \\ &= O\left(\delta_n^{-\alpha_2}|\xi_{i_0}|^3 \left(\sum_{i=1}^k |a_{i,i_0}|\right)^3\right) = O\big(\delta_n^{-\alpha_2}|\xi_{i_0}|^3 a_{i_0}^3\big). \end{aligned} \tag{69}$$

Note that as in (67), $\hat{L}_2 = O(\delta_n^{-\alpha_2}|\xi_{i_0}|^2 a_{i_0}^2)$. Thus,

$$\mathrm{err}_2 = \mathrm{err}_1 - \frac{\hat{L}_2}{2} = O\big(\delta_n^{-\alpha_2}|\xi_{i_0}|^2 a_{i_0}^2\big). \tag{70}$$

Using (69) and (70), we obtain

$$|\mathrm{err}_2| = O\big(\delta_n^{-\alpha_2}\big)\min\big\{|\xi_{i_0}|^2 a_{i_0}^2, |\xi_{i_0}|^3 a_{i_0}^3\big\} = O\big(\delta_n^{-\alpha_2}|\xi_{i_0}|^{2+\varepsilon} a_{i_0}^{2+\varepsilon}\big). \tag{71}$$

The expression (68) also holds for $\tilde{\xi}$ in place of ξ; we denote the error term here by $\widetilde{\mathrm{err}_2}$. By the same reasoning, we can show that $\widetilde{\mathrm{err}_2}$ satisfies (71).

Take the expectation (with respect to ξ_{i_0}) of the right-hand side of (68) and subtract from it the expectation of the corresponding formula (with respect to $\tilde{\xi}_{i_0}$). By Condition C1, ξ_{i_0} and $\tilde{\xi}_{i_0}$ have matching first and second moments, and so the expectations of $\hat{L}_j$ $(j = 1,2)$ from the two formulae cancel each other out. Furthermore, $\hat{L}(0)$ is the same in both formulae. Thus, the only thing remaining after the subtraction are the error terms. Therefore,

$$\begin{aligned} d_{i_0} \le \left|\mathbf{E}_{\xi_{i_0}} \mathrm{err}_2\right| + \left|\mathbf{E}_{\tilde{\xi}_{i_0}} \widetilde{\mathrm{err}_2}\right| &= O(1)\tilde{C}\delta_n^{-\alpha_2} a_{i_0}^{2+\varepsilon}\left(\mathbf{E}|\xi_{i_0}|^{2+\varepsilon} + \mathbf{E}|\tilde{\xi}_{i_0}|^{2+\varepsilon}\right) \\ &= O\left(\delta_n^{-\alpha_2} a_{i_0}^{2+\varepsilon}\right). \end{aligned}$$

Taking expectation with respect to the other variables (which we have conditioned on so far), we obtain $I_{i_0} = O(\delta_n^{-\alpha_2} a_{i_0}^{2+\varepsilon})$ for all $N_0 \le i_0 \le n$.

Now we treat the first few indices $0 \le i_0 < N_0$, where ξ_{i_0} may have non-zero mean. Instead of using (66) and (68), we use the mean value theorem to get the rough bound

$$\hat{L}\left(a_{1,i_0}\xi_{i_0}, \dots, a_{k,i_0}\xi_{i_0}\right) = \hat{L}(0) + O\left(k\left\|\nabla \hat{L}\right\|_\infty |\xi_{i_0}| \sum_{i=1}^k |a_{i,i_0}|\right), \tag{72}$$

which by the same arguments as above gives $I_{i_0} = O(\delta_n^{-\alpha_2} a_{i_0})$.

Since we assume $\mathbf{E}\tilde{\xi}_i = 0$ for all $1 \le i \le n$, Condition C1 implies that $|\mathbf{E}\xi_{i_0}| = O(1)$. But as $\operatorname{Var}\xi_{i_0} = 1$, it follows that $\mathbf{E}|\xi_{i_0}| = O(1)$.

As k is constant and $\left\|\nabla \hat{L}\right\|_\infty \le \delta_n^{-\alpha_2}$, we have, from (72), that

$$\begin{aligned} d_{i_0} &= O\left(k\|\nabla \hat{L}\|_\infty \sum_{i=1}^k |a_{i,i_0}|\right)\left(\mathbf{E}|\xi_{i_0}| + \mathbf{E}|\tilde{\xi}_{i_0}|\right) \\ &= O\left(\delta_n^{-\alpha_2} \sum_{i=1}^k |a_{i,i_0}|\right) = O\left(\delta_n^{-\alpha_2} \sum_{i=1}^k |a_{i,i_0}|^2\right)^{1/2} = O\left(\delta_n^{-\alpha_2} a_{i_0}\right). \end{aligned}$$

Notice that by Condition C2(4), $a_{i_0} = O(\sqrt{k}\delta_n^{\alpha_1}) = O(\delta_n^{\alpha_1})$ for all i. Furthermore, by the definition $\sum_{i=N_0}^n a_{i_0}^2 = k = O(1)$. Thus, we have

$$I = O\left(\delta_n^{-\alpha_2} \sum_{i_0=0}^n a_{i_0}^{2+\varepsilon} + \delta_n^{-\alpha_2} \sum_{i_0=0}^{N_0} a_{i_0}\right) = O\left(\delta_n^{\alpha_1\varepsilon - \alpha_2}\right) = O\left(\delta_n^{\alpha_2}\right),$$

wherein the last step we used the fact that α_2 was set much smaller than α_1. □

Proof of Lemma 8.3. Let α_2 be the constant in Lemma 15.1 and set $\alpha_0 := \frac{\alpha_2}{10}$. Let

$$\bar{K}(w_1,\dots,w_k) := K\left(w_1 + \frac{1}{2}\log\left|V(z_1)\right|,\dots,w_k + \frac{1}{2}\log\left|V(z_k)\right|\right)$$

where we recall that $V(z_j) := \sum_{i=N_0}^{n} |\phi_i(z_j)|^2$. We have $\left\|\nabla^a \bar{K}\right\|_\infty \le \delta_n^{-\alpha_0}$ for all $0 \le \alpha \le 3$; we aim to show

$$\begin{aligned}(73)\quad &\left|\mathbf{E}\bar{K}\left(\log\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\dots,\log\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\right)\right.\\ &\left.\quad -\mathbf{E}\bar{K}\left(\log\frac{|\tilde{F}_n(z_1)|}{\sqrt{V(z_1)}},\dots,\log\frac{|\tilde{F}_n(z_k)|}{\sqrt{V(z_k)}}\right)\right| = O\left(\delta_n^{\alpha_0}\right).\end{aligned}$$

For $M := \log(\delta_n^{-3\alpha_0})$, define

$$\Omega_1 := \left\{(w_1,\dots,w_k) \in \mathbb{R}^k : \min_{i=1,\dots,k} w_i < -M\right\}$$

and

$$\Omega_2 := \left\{(w_1,\dots,w_k) \in \mathbb{R}^k : \min_{i=1,\dots,k} w_i > -M-1\right\}.$$

By considering the real and imaginary parts of $\bar{K}$ separately, we can assume that $\bar{K} : \mathbb{R}^k \to \mathbb{R}$.

Let $\psi : \mathbb{R}^k \to [0,1]$ be a smooth function supported in Ω_2 such that $\psi = 1$ on the complement of Ω_1 and $\|\nabla^a \psi\|_\infty = O(1)$ for all $0 \le a \le 3$. As $M \ge 1$, it is easy to see that such a function exists. In particular, one can define $\psi(x_1,\dots,x_k) = \rho(x_1)\dots\rho(x_k)$ where ρ is a smooth function satisfying the corresponding properties on $\mathbb{R}$.

Let $\phi := 1-\psi$, $K_1 := \bar{K}\phi$, and $K_2 := \bar{K}\psi$. Then by the definition $\bar{K} = K_1 + K_2$. Furthermore, both K_1, K_2 are smooth functions with $\operatorname{supp} K_1 \subset \bar{\Omega}_1, \operatorname{supp} K_2 \subset \bar{\Omega}_2$ and $\|\nabla^a K_i\|_\infty = O(\delta_n^{-\alpha_0})$ for $i = 1,2$ and $0 \le a \le 3$.

We now show that the contribution from K_1 towards the right-hand side of (73) is negligible. Notice that

$$\|K_1\|_\infty \le \left\|\bar{K}\right\|_\infty \le C'\delta_n^{-\alpha_0}.$$

This leads to setting $H_1(w_1,\dots,w_k) = C'\delta_n^{-\alpha_0}\phi(\log|w_1|,\dots,\log|w_k|)$. The function H_1 is a smooth function on $\mathbb{R}^k$ with the following properties:

- $|K_1(\log|w_1|,\dots,\log|w_k|)| \le H_1(w_1,\dots,w_k)$,
- $\operatorname{supp}(H_1) \subset \{(w_1,\dots,w_k) \in \mathbb{R}^k : \min_{i=1,\dots,k}|w_i| \le e^{-M}\}$,
- $\|\nabla^a H_1\|_\infty = O(\delta_n^{-10\alpha_0}) = O(\delta^{-\alpha_2})$ for all $0 \le a \le 3$.

Remark 15.3. To verify the last property, notice that the support of H_1 is $\{(x,y) : |x| \le e^{-M} \text{ or } |y| \le e^{-M}\}$. Moreover, H_1 is a constant $C'\delta_n^{-\alpha_0}$ in the set

$\{(x,y) : |x| \le e^{-M-1}$ or $|y| \le e^{-M-1}\}$ (because $\phi = 1$ on the complement of Ω_2). So we only need to consider the derivatives of H_1 in the set $\{(x,y) : |x| \le b$ or $|y| \le e^{-M}\} \cap \{|x| \ge e^{-M-1}, |y| \ge e^{-M-1}\}$. On that set, x^{-1} and y^{-1} are bounded from above by e^{M+1}, which is significantly smaller than the bound. (We define α_0 and M with foresight so the claimed bound holds, with room to spare.)

Applying Lemma 15.1, we obtain

$$\begin{aligned}\mathbf{E}\left|K_1\left(\log\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\ldots,\log\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\right)\right| &\le \mathbf{E}H_1\left(\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\ldots,\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\right)\\ &\le \mathbf{E}H_1\left(\frac{|\tilde{F}_n(z_1)|}{\sqrt{V(z_1)}},\ldots,\frac{|\tilde{F}_n(z_k)|}{\sqrt{V(z_k)}}\right)\\ &\quad + C'\delta_n^{\alpha_0}.\end{aligned}$$

Since $H_1(w_1,\ldots,w_k) = 0$ if $(\log|w_1|,\ldots,\log|w_k|) \notin \Omega_1$ and since the variables $\tilde{\xi}_i$ are Gaussian, we have

$$\begin{aligned}&\mathbf{E}H_1\left(\frac{|\tilde{F}_n(z_1)|}{\sqrt{V(z_1)}},\ldots,\frac{|\tilde{F}_n(z_k)|}{\sqrt{V(z_k)}}\right)\\ &\quad\le C'\delta_n^{-\alpha_0}\mathbf{P}\left(\exists i \in \{1,\ldots,k\} : \frac{|\tilde{F}_n(z_i)|}{\sqrt{V(z_i)}} \le e^{-M} = \delta_n^{3\alpha_0}\right)\\ &\quad\le C'\delta_n^{-\alpha_0}k\delta_n^{3\alpha_0} = O\big(\delta_n^{\alpha_0}\big).\end{aligned}$$

Thus, $\mathbf{E}\big|K_1\big(\log\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\ldots,\log\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\big)\big| \le C'\delta_n^{\alpha_0}$. The same bound holds with F_n replaced by $\tilde{F}_n$. To conclude the proof, we need to show that

$$\begin{aligned}&\left|\mathbf{E}K_2\left(\log\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\ldots,\log\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\right)\right.\\ &\qquad\left. - \mathbf{E}K_2\left(\log\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\ldots,\log\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\right)\right| = O\big(\delta_n^{\alpha_0}\big).\end{aligned}$$

Define $H_2(w_1,\ldots,w_k) := K_2(\log|w_1|,\ldots,\log|w_2|)$. Since $\operatorname{supp} K_2 \subset \bar{\Omega}_2$,

$$\begin{aligned}\operatorname{supp} H_2 &\subset \big\{(w_1,\ldots,w_k) : \log|w_i| \ge -M-1,\ \forall i\big\}\\ &= \big\{(w_1,\ldots,w_k) : |w_i| \ge C'\delta_n^{3\alpha_0},\ \forall i\big\}.\end{aligned}$$

Thus, H_2 is well defined and smooth on $\mathbb{R}^k$. Furthermore, by the definition of H_2, it is not hard to check that $\|\nabla^a H_2\|_\infty = O(\delta_n^{-10\alpha_0})$ for all $0 \le a \le 3$; see Remark 15.3.

Finally, by applying Lemma 15.1, we obtain

$$\begin{aligned}&\Bigg|\mathbf{E}K_2\Big(\log\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\dots,\log\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\Big)\\&\quad-\mathbf{E}K_2\Big(\log\frac{|\tilde F_n(z_1)|}{\sqrt{V(z_1)}},\dots,\log\frac{|\tilde F_n(z_k)|}{\sqrt{V(z_k)}}\Big)\Bigg|\\&=\Bigg|\mathbf{E}H_2\Big(\frac{|F_n(z_1)|}{\sqrt{V(z_1)}},\dots,\frac{|F_n(z_k)|}{\sqrt{V(z_k)}}\Big)\\&\quad-\mathbf{E}H_2\Big(\frac{|\tilde F_n(z_1)|}{\sqrt{V(z_1)}},\dots,\frac{|\tilde F_n(z_k)|}{\sqrt{V(z_k)}}\Big)\Bigg|=O\big(\delta_n^{\alpha_0}\big).\end{aligned}$$

This completes the proof. □

15.4. Proof of Lemma 9.3. By rescaling, we can assume that $a=n^l$. Thus, we need to estimate $\sup_Z \mathbf{P}\big(\big|\sum_{j=1}^n a_j\varepsilon_j - Z\big|\le 1\big)$.

By Esséen's inequality [16] (see also [57, Lemma 7.17]), there is an absolute constant c such that for any real number Z,

$$\mathbf{P}\left(\left|\sum_{j=1}^n a_j\varepsilon_j - Z\right|\le 1\right)\le c\int_{-1/2}^{1/2}|\phi(t)|dt \tag{74}$$

where

$$\phi(t)=\mathbf{E}\exp\left(i2\pi t\sum_{j=1}^n a_j\varepsilon_j\right)=\prod_{j=1}^n\mathbf{E}\exp\big(i2\pi t a_j\varepsilon_j\big)=\prod_{j=1}^n\cos(2\pi a_j t).$$

For every $x\in\mathbb{R}$, let $\|x\|_{\mathbb{R}/\mathbb{Z}}:=\min\{|x-N|:N\in\mathbb{Z}\}$ be the distance from x to the set of integers. In the following lemma, we gather a few simple (and well-known) facts concerning sin and cos, whose proof is left as an exercise.

LEMMA 15.4. *We have*

- $\sin\theta\ge 2\theta/\pi$ *for all* $\theta\in[0,\pi/2]$;
- $|\cos x|\le 1-2\|x/\pi\|^2_{\mathbb{R}/\mathbb{Z}}\le\exp(-2\|x/\pi\|^2_{\mathbb{R}/\mathbb{Z}})$ *for all* $x\in\mathbb{R}$;
- $\cos(2x)\ge 1-2\pi^2\|x/\pi\|^2_{\mathbb{R}/\mathbb{Z}}$ *for all* $x\in\mathbb{R}$;
- *There is a constant* $c>0$ *such that for all* $T\ge 1$,

$$\max\left\{\left|\int_0^1\sin Tx\,dx\right|,\left|\int_0^1\cos Tx\,dx\right|\right\}\le c/T.$$

By (74) and Fubini's Theorem,

$$\mathbf{P}\left(\left|\sum_{j=1}^{n} a_j\varepsilon_j - Z\right| \le 1\right) \le c\int_{-1/2}^{1/2} \exp\left(-2\sum_{j=1}^{n}\|2a_jt\|_{\mathbb{R}/\mathbb{Z}}^2\right)dt \tag{75}$$
$$= 2c\int_0^\infty |A_x|e^{-2x}dx,$$

where $A_x := \{t \in [-1/2, 1/2] : \sum_{j=1}^{n}\|2a_jt\|_{\mathbb{R}/\mathbb{Z}}^2 \le x\}$ and $|A_x|$ denotes the Lebesgue measure of A_x. We break the last integral in (75) into two parts, $\int_0^{n/4\pi^2}|A_x|e^{-2x}dx$ and $\int_{n/4\pi^2}^\infty |A_x|e^{-2x}dx$. Since $|A_x| \le 1$ for all x,

$$\int_{n/4\pi^2}^\infty |A_x|e^{-2x}dx = e^{-\Omega(n)} = o\big(n^{-l}\big)$$

for any fixed l. Thus, this part is negligible and it remains to show

$$\int_0^{n/4\pi^2}|A_x|e^{-2x}dx = O(n^{-l}). \tag{76}$$

Let us now bound the measure of the set $A_{n/4\pi^2}$. By Lemma 15.4,

$$A_{n/4\pi^2} \subset A := \left\{t \in [-1/2, 1/2] : \sum_{j=1}^{n}\cos(4\pi a_jt) \ge n/2\right\}.$$

To bound $|A|$, we first notice that

$$\int_{-1/2}^{1/2}\left(\sum_{j=1}^{n}\cos(4\pi a_jt)\right)^{2l}dt \le \int_{-1/2}^{1/2}\left(\sum_{j=1}^{n}\big(e^{i4\pi a_jt} + e^{-i4\pi a_jt}\big)\right)^{2l}dt$$
$$= \sum_{s_1,\dots,s_{2l}=\pm 1}\ \sum_{j_1,\dots,j_{2l}\le n}\int_{-1/2}^{1/2} e^{i4\pi t\sum_{h=1}^{2l}s_ha_{j_h}}dt.$$

Recall the hypothesis of the lemma that for any two different multi-sets $\{i_1,\dots,i_{l'}\}$ and $\{j_1,\dots,j_{l''}\}$ where $l'+l'' \le 2l$, it holds that $|a_{i_1}+\dots+a_{i_{l'}}-a_{j_1}-\dots-a_{j_{l''}}| \ge a = n^l$. Thus, for each $s_1,\dots,s_{2l} = \pm 1$ and $j_1,\dots,j_{2l} \le n$, consider the multi-sets $S_1 = \{j_h : s_h = 1\}$ and $S_2 = \{j_h : s_h = -1\}$. If $S_1 \ne S_2$ then $|\sum_{h=1}^{2l}s_ha_{j_h}| \ge n^l$. In this case, the corresponding term in the above double sum is of the form $\int_{-1/2}^{1/2}e^{itT}dt$ for some $|T| \ge 2n^l$. By Lemma 15.4, we have

$$\int_{-1/2}^{1/2} e^{i4\pi t\sum_{h=1}^{2l}s_ha_{j_h}}dt = O\big(n^{-l}\big), \quad \text{if } S_1 \ne S_2.$$

If $S_1 = S_2$, then $|a_{i_1}+\dots+a_{i_{l'}}-a_{j_1}-\dots-a_{j_{l''}}| = 0$ and the corresponding integral is 1. The number of terms in the double sum with $S_1 = S_2$ is at most $2^{2l}n^l = O(n^l)$

while the total number of terms is at most $2^{2l}n^{2l} = O(n^{2l})$. Putting these cases together, we obtain

$$\int_{-1/2}^{1/2} \left(\sum_{j=1}^{n} \cos\left(4\pi a_j t\right) \right)^{2l} dt = O\left(n^l + n^{2l} n^{-l}\right) = O\left(n^l\right).$$

Hence, $|A| = O(n^{-l})$ by Markov's inequality. This implies $|A_{n/4\pi^2}| = O(n^{-l})$, which, in turn, yields (76), completing the proof.

15.5. Proof of the second Jensen's inequality (8). By setting $g(w) = f(R(w+z))$ and prove the corresponding inequality for g, it suffices to assume that $z = 0$ and $R = 1$. Let $a_1, \ldots, a_N$ be the zeros of f in $\bar{B}(0,r)$. For each a inside the unit disk D, consider the map

$$T_a(w) = \frac{w-a}{\bar{a}w - 1}.$$

For $|a| \le r$ and $|w| \le r$, one can show by algebraic manipulation that

$$|T_a(w)| \le \frac{2r}{1+r^2} < 1.$$

Moreover, for all $|a| < 1$ and $|w| = 1$, we have

$$|T_a(w)| = |\bar{w}| \left| \frac{w-a}{\bar{a}w-1} \right| = \left| \frac{1 - a\bar{w}}{\bar{a}w - 1} \right| = 1.$$

Let $h(w) = \frac{f(w)}{\prod_{k=1}^{N} T_{a_k}(w)}$. Then h is an analytic function on D. By maximum principle, we have for every $w_0 \in rD$,

$$\begin{aligned} \frac{\left|f(w_0)\right| \left(1+r^2\right)^N}{(2r)^N} &\le \max_{w \in rD} \left|h(w)\right| \le \max_{w \in D} \left|h(w)\right| \\ &= \max_{w \in \partial D} \left|h(w)\right| = \max_{w \in \partial D} \left|f(w)\right| = M. \end{aligned}$$

Thus, $N \le \frac{\log \frac{M}{|f(w_0)|}}{\log \frac{1+r^2}{2r}}$ for all $w_0 \in rD$, completing the proof.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, IL 61801
E-mail: onguyen@illinois.edu

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CT 06520
E-mail: van.vu@yale.edu

REFERENCES

[1] J. Angst and G. Poly, A weak Cramér condition and application to Edgeworth expansions, *Electron. J. Probab.* **22** (2017), Paper No. 59.

[2] J.-M. Azaïs, F. Dalmao, J. León, I. Nourdin, and G. Poly, Local universality of the number of zeros of random trigonometric polynomials with continuous coefficients, preprint, https://arxiv.org/abs/1512.05583.

[3] A. T. Bharucha-Reid and M. Sambandham, *Random Polynomials*, *Probab. Math. Statist.*, Academic Press, Orlando, FL, 1986.

[4] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation*, *Encyclopedia Math. Appl.*, vol. 27, Cambridge University Press, Cambridge, 1989.

[5] P. Bleher and X. Di, Correlations between zeros of a random polynomial, *J. Statist. Phys.* **88** (1997), no. 1-2, 269–305.

[6] ———, Correlations between zeros of non-Gaussian random polynomials, *Int. Math. Res. Not. IMRN* **2004** (2004), no. 46, 2443–2484.

[7] A. Bloch and G. Pólya, On the roots of certain algebraic equations, *Proc. London Math. Soc. (2)* **33** (1931), no. 1, 102–114.

[8] M.-C. Chang, H. Nguyen, O. Nguyen, and V. Vu, Random eigenfunctions on flat tori: universality for the number of intersections, *Int. Math. Res. Not. IMRN* **2020** (2020), no. 24, 9933–9973.

[9] M. Das, The average number of real zeros of a random trigonometric polynomial, *Proc. Cambridge Philos. Soc.* **64** (1968), 721–729.

[10] Y. Do, H. Nguyen, and V. Vu, Real roots of random polynomials: expectation and repulsion, *Proc. Lond. Math. Soc. (3)* **111** (2015), no. 6, 1231–1260.

[11] Y. Do, O. Nguyen, and V. Vu, Roots of random polynomials with coefficients of polynomial growth, *Ann. Probab.* **46** (2018), no. 5, 2407–2494.

[12] ———, Random orthonormal polynomials: local universality and expected number of real roots, preprint, https://arxiv.org/abs/2012.10850.

[13] J. E. A. Dunnage, The number of real zeros of a random trigonometric polynomial, *Proc. London Math. Soc. (3)* **16** (1966), 53–84.

[14] A. Edelman and E. Kostlan, How many zeros of a random polynomial are real?, *Bull. Amer. Math. Soc. (N.S.)* **32** (1995), no. 1, 1–37.

[15] P. Erdös and A. C. Offord, On the number of real roots of a random algebraic equation, *Proc. London Math. Soc. (3)* **6** (1956), 139–160.

[16] C.-G. Esseen, On the Kolmogorov-Rogozin inequality for the concentration function, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **5** (1966), 210–216.

[17] K. Farahmand, *Topics in Random Polynomials*, *Pitman Res. Notes Math. Ser.*, vol. 393, Longman, Harlow, 1998.

[18] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. I*, 3rd ed., John Wiley & Sons, New York, 1968.

[19] H. Flasche, Expected number of real roots of random trigonometric polynomials, *Stochastic Process. Appl.* **127** (2017), no. 12, 3928–3942.

[20] H. Flasche and Z. Kabluchko, Expected number of real zeroes of random Taylor series, *Commun. Contemp. Math.* **22** (2020), no. 7, 1950059.

[21] ———, Real zeroes of random analytic functions associated with geometries of constant curvature, *J. Theoret. Probab.* **33** (2020), no. 1, 103–133.

[22] G. B. Folland, *Real Analysis. Modern Techniques and Their Applications*, 2nd ed., *Pure Appl. Math. (New York)*, John Wiley & Sons, New York, 1999.

[23] P. J. Forrester and G. Honner, Exact statistical properties of the zeros of complex random polynomials, *J. Phys. A* **32** (1999), no. 16, 2961–2981.

[24] F. Götze, D. Kaliada, and D. Zaporozhets, Correlation functions of real zeros of random polynomials, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **454** (2016), 102–111.

[25] G. Halász, Estimates for the concentration function of combinatorial number theory and probability, *Period. Math. Hungar.* **8** (1977), no. 3-4, 197–211.

[26] J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág, *Zeros of Gaussian analytic functions and determinantal point processes*, *Univ. Lecture Ser.*, vol. 51, Amer. Math. Soc., Providence, RI, 2009.

[27] I. A. Ibragimov and N. B. Maslova, The average number of zeros of random polynomials, *Vestnik Leningrad. Univ.* **23** (1968), no. 19, 171–172.

[28] ———, On the expected number of real zeros of random polynomials. I: Coefficients with zero means, *Theory Probab. Appl.* **16** (1971), 228–248.

[29] A. Iksanov, Z. Kabluchko, and A. Marynych, Local universality for real roots of random trigonometric polynomials, *Electron. J. Probab.* **21** (2016), Paper No. 63.

[30] Z. Kabluchko and D. Zaporozhets, Asymptotic distribution of complex zeros of random analytic functions, *Ann. Probab.* **42** (2014), no. 4, 1374–1395.

[31] ———, Universality for zeros of random analytic functions, preprint, https://arxiv.org/abs/1205.5355.

[32] M. Kac, On the average number of real roots of a random algebraic equation, *Bull. Amer. Math. Soc.* **49** (1943), 314–320.

[33] ———, On the average number of real roots of a random algebraic equation (II), *Proc. London Math. Soc. (2)* **50** (1948), 390–408.

[34] J.-P. Kahane, *Some Random Series of Functions*, 2nd ed., *Cambridge Stud. Adv. Math.*, vol. 5, Cambridge University Press, Cambridge, 1985.

[35] A. Ledoan, M. Merkli, and S. Starr, A universality property of Gaussian analytic functions, *J. Theoret. Probab.* **25** (2012), no. 2, 496–504.

[36] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation (III), *Rec. Math. [Mat. Sbornik] N.S.* **12(54)** (1943), 277–286.

[37] ———, On the distribution of the zeros and α-values of a random integral function (I), *J. London Math. Soc.* **20** (1945), 130–136.

[38] ———, On the distribution of zeros and α-values of a random integral function (II), *Ann. of Math. (2)* **49** (1948), 885–952.

[39] F. Nazarov, A. Nishry, and M. Sodin, Log-integrability of Rademacher Fourier series, with applications to random analytic functions, *St. Petersburg Math. J.* **25** (2014), no. 3, 467–494.

[40] F. L. Nazarov, Local estimates of exponential polynomials and their applications to inequalities of the uncertainty principle type, *St. Petersburg Math. J.* **5** (1994), 663–717.

[41] H. Nguyen, O. Nguyen, and V. Vu, On the number of real roots of random polynomials, *Commun. Contemp. Math.* **18** (2016), no. 4, 1550052, 17.

[42] H. Nguyen and V. Vu, Optimal inverse Littlewood-Offord theorems, *Adv. Math.* **226** (2011), no. 6, 5298–5319.

[43] ———, Small ball probability, inverse theorems, and applications, *Erdös Centennial*, *Bolyai Soc. Math. Stud.*, vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 409–463.

[44] I. E. Pritsker, Zero distribution of random polynomials, *J. Anal. Math.* **134** (2018), no. 2, 719–745.

[45] I. E. Pritsker and A. M. Yeager, Zeros of polynomials with random coefficients, *J. Approx. Theory* **189** (2015), 88–100.

[46] T. Prosen, Parametric statistics of zeros of Husimi representations of quantum chaotic eigenstates and random polynomials, *J. Phys. A* **29** (1996), no. 17, 5429–5440.

[47] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1987.

[48] M. Sambandham, On the number of real zeros of a random trigonometric polynomial, *Trans. Amer. Math. Soc.* **238** (1978), 57–70.

[49] G. Schehr and S. N. Majumdar, Real roots of random polynomials and zero crossing properties of diffusion equation, *J. Stat. Phys.* **132** (2008), no. 2, 235–273.

[50] Y. Shalom and T. Tao, A finitary version of Gromov's polynomial growth theorem, *Geom. Funct. Anal.* **20** (2010), no. 6, 1502–1547.

[51] M. Sodin, Zeroes of Gaussian analytic functions, *European Congress of Mathematics*, Eur. Math. Soc., Zürich, 2005, pp. 445–458.

[52] M. Sodin and B. Tsirelson, Random complex zeroes. I. Asymptotic normality, *Israel J. Math.* **144** (2004), 125–149.

[53] K. Söze, Real zeroes of random polynomials, I. Flip-invariance, Turán's lemma, and the Newton-Hadamard polygon, *Israel J. Math.* **220** (2017), no. 2, 817–836.

[54] ———, Real zeroes of random polynomials, II. Descartes' rule of signs and anti-concentration on the symmetric group, *Israel J. Math.* **220** (2017), no. 2, 837–872.

[55] S. Starr, Universality of correlations for random analytic functions, *Entropy and the Quantum II*, *Contemp. Math.*, vol. 552, Amer. Math. Soc., Providence, RI, 2011, pp. 135–144.

[56] D. C. Stevens, The average number of real zeros of a random polynomial, *Comm. Pure Appl. Math.* **22** (1969), 457–477.

[57] T. Tao and V. Vu, *Additive Combinatorics*, *Cambridge Stud. Adv. Math.*, vol. 105, Cambridge University Press, Cambridge, 2006.

[58] ———, Local universality of zeroes of random polynomials, *Int. Math. Res. Not. IMRN* **2015** (2015), no. 13, 5053–5139.

[59] P. Turán, *Eine neue Methode in der Analysis und deren Anwendungen*, Akadémiai Kiadó, Budapest, 1953.

[60] J. E. Wilkins Jr., Mean number of real zeros of a random trigonometric polynomial, *Proc. Amer. Math. Soc.* **111** (1991), no. 3, 851–863.

[61] S. Zelditch, From random polynomials to symplectic geometry, *XIIIth International Congress on Mathematical Physics (London, 2000)*, Int. Press, Boston, MA, 2001, pp. 367–376.