The Annals of Probability
2018, Vol. 46, No. 5, 2407–2494
https://doi.org/10.1214/17-AOP1219

ROOTS OF RANDOM POLYNOMIALS WITH COEFFICIENTS OF POLYNOMIAL GROWTH

BY YEN DO*,[1], OANH NGUYEN†,[2] AND VAN VU†,[2]

University of Virginia and Yale University*†

In this paper, we prove optimal local universality for roots of random polynomials with arbitrary coefficients of polynomial growth. As an application, we derive, for the first time, sharp estimates for the number of real roots of these polynomials, even when the coefficients are not explicit. Our results also hold for series; in particular, we prove local universality for random hyperbolic series.

CONTENTS

1. A motivation: Real roots of random polynomials . . . 2408
   1.1. Description of the new results for coefficients with zero means . . . 2411
   1.2. Polynomials with coefficients having nonzero means . . . 2415
   1.3. Outline of the paper . . . 2416
2. Correlation functions and universality . . . 2416
   2.1. Universality . . . 2418
   2.2. Coefficients with polynomial growth . . . 2419
   2.3. Complex local universality for polynomials . . . 2419
   2.4. Real local universality . . . 2421
   2.5. Local universality for series . . . 2423
3. Sketch of the proof and the main technical ideas . . . 2425
4. Proof of complex local universality for polynomials . . . 2430
   4.1. The event $\mathcal{T}$ occurs with high probability . . . 2431
   4.2. Approximation of integrals by finite sums . . . 2437
   4.3. Log comparability . . . 2440
   4.4. On the tail event $\mathcal{T}^c$ . . . 2441
   4.5. Finishing . . . 2450
5. Proof of real local universality . . . 2451
6. Proof of Lemma 2.5 and Corollary 2.6 . . . 2460
7. Proof of complex local universality for series . . . 2465
8. Proof of Theorem 2.8, part I: Reduction to the case $M = m = 1$ . . . 2468
9. Proof of Theorem 2.8, part II: Estimates for the density function near ± 1 . . . 2469
   9.1. Convergence of ρ_n . . . 2470
   9.2. Blowup nature of ρ_∞ . . . 2471

Received May 2016; revised July 2017.

[1]Supported in part by NSF Grants DMS-1201456 and DMS-1521293.

[2]Supported in part by NSF Grant DMS-1307797 and AFORS Grant FA9550-12-1-0083.

MSC2010 subject classifications. 60K35.

Key words and phrases. Random polynomials, real roots, complex roots, correlation, arbitrary coefficients, universality.

10. Proof of Theorem 2.8, part III: Counting real zeros near ± 1 . . . 2471
10.1. Proof of Lemma 10.1 . . . 2472
10.2. Proof of Lemma 10.2 . . . 2475
11. Proof of Theorem 2.9 . . . 2481
Appendix . . . 2487
Acknowledgments . . . 2492
References . . . 2492

1. A motivation: Real roots of random polynomials. Let us start by describing a natural and famous problem which serves as the motivation of our studies, the main results of which will be discussed in Section 2.

Finding real roots of a high degree polynomial is among the most basic problems in mathematics. From the algebraic point of view, it is classical that for most polynomials of degree at least 5, the roots cannot be computed in radicals, thanks to the fundamental works of Abel–Ruffini and Galois. There has been a huge amount of results on the number of real roots and also their locations using information from the coefficients (for instance, one of the earliest results is Descartes' classical theorem concerning sign sequences); however, most results are often sharp for certain special classes of polynomials, but poor in many others.

It is natural and important to consider the root problem from the statistical point of view. What can we say about a *typical* (*i.e.*, *random*) polynomial? Already in the seventeenth century, Waring considered random cubic polynomials and concluded that the probability of having three real roots is at most $2/3$. This effort was discussed by Toddhunter in [40], one of the earliest books in probability theory, which also reported a similar effort made by Sylvester. However, the distribution of the polynomials was not explicitly defined at the time.

In the last hundred years, random polynomials have attracted the attention of many generations of mathematicians, with most efforts directed to the following model:

$$P_{n,\xi}(x) := c_n\xi_n x^n + \cdots + c_1\xi_1 x + c_0\xi_0 x^0,$$

where ξ_i are i.i.d. copies of a random variables ξ with zero mean and unit variance, and c_i are deterministic coefficients which may depend on both n and i. Different definitions of c_i give rise to different classes of random polynomials, which have different behaviors. When $c_i = 1$ for all i, the polynomial $P_{n,\xi}$ is often referred to as the Kac polynomial. Even for this special case, the literature is very rich (see [3, 12] for surveys). In the next few paragraphs, we will discuss few seminal results which directly motivate our research.

The first modern work on random polynomials was due to Bloch and Polya in 1932 [5], who considered the Kac polynomial with ξ being Rademacher [namely $\mathbf{P}(\xi = 1) = \mathbf{P}(\xi = -1) = 1/2$], and showed that with high probability

$$N_{n,\xi} = O(\sqrt{n}),$$

where $N_{n,\xi}$ denotes the number of real roots of the Kac polynomial associated with the random variable ξ. Their key idea is simple and beautiful. Notice that if we apply Descartes' rule of signs for P_n, one could only obtain the trivial bound $O(n)$ for $N_{n,\xi}$ as the typical number of sign changes is around $n/2$. Bloch and Polya's idea is to apply Descartes rule for $P_n Q$, where Q is a deterministic polynomial which does not have any real positive roots. By choosing Q properly, they reduced the number of sign changes significantly.

Next came the ground breaking series of papers by Littlewood and Offord [20–22] in the early 1940s, which, to the surprise of many mathematicians of their time, showed that $N_{n,\xi}$ is typically polylogarithmic in n.

THEOREM 1.1 (Littlewood–Offord). *For ξ being Rademacher, Gaussian or uniform on $[-1,1]$,*

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

with probability $1-o(1)$.

Littlewood–Offord's papers and later works of Offord [20–22] lay the foundation for the theory of random functions, which is an important part of modern probability and analysis; see, for instance, [28, 33].

During more or less the same time, Kac [18] discovered his famous formula for the density function $\rho(t)$ of $N_{n,\xi}$

$$\rho(t) = \int_{-\infty}^{\infty} |y| p(t,0,y)\,dy, \tag{1.1}$$

where $p(t,x,y)$ is the joint probability density for $P_{n,\xi}(t)=x$ and the derivative $P'_{n,\xi}(t)=y$.

Consequently,

$$\mathbf{E}N_{n,\xi} = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} |y| p(t,0,y)\,dy. \tag{1.2}$$

In the Gaussian case (ξ is Gaussian), the joint distribution of $P_{n,\xi}(t)$ and $P'_{n,\xi}(t)$ can be explicitly computed. Using this fact, Kac showed in [18] that

$$\mathbf{E}N_{n,\mathrm{Gauss}} = \frac{1}{\pi}\int_{-\infty}^{\infty} \sqrt{\frac{1}{(t^2-1)^2} + \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}}\,dt = \left(\frac{2}{\pi}+o(1)\right)\log n. \tag{1.3}$$

A more careful evaluation by Wilkins [41] and also Edelman and Kostlan [10] gives

$$\mathbf{E}N_{n,\mathrm{Gauss}} = \frac{2}{\pi}\log n + C + o(1), \tag{1.4}$$

where C is an explicit constant. As a matter of fact, Wilkins [41] computed all terms in the Taylor expansion of the integration in (1.3).

In his original paper [18], 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 to be not the case, as the right-hand side of (1.2) is often hard to compute, especially when ξ is discrete (Rademacher for instance). Technically, the joint distribution of $P_{n,\xi}(t)$ and $P'_{n,\xi}(t)$ is easy to determine in the Gaussian case, thanks to special properties of the Gaussian distribution, but can pose a great challenge in the general one. Kac admitted this in a later paper [19], in which he managed to push his method to treat the case ξ being uniform in $[-1, 1]$. A further extension was made by Stevens [37], who evaluated Kac's formula for a large class of ξ having continuous and smooth distributions with certain regularity properties (see [37], page 457, for details).

The treatment of $\mathbf{E}N_{n,\xi}$ for discrete random variables ξ required considerable effort. More than 10 years after Kac's paper [18], Erdős and Offord [11] found a completely new approach to handle the Rademacher case, proving that with probability $1 - o(\frac{1}{\sqrt{\log\log n}})$

$$N_{n,\xi} = \frac{2}{\pi}\log n + o\big(\log^{2/3} n \log\log n\big). \tag{1.5}$$

The argument of Erdős and Offord is combinatorial and very delicate, even by today's standard. Their main idea is to approximate the number of roots by the number of sign changes in $P_{n,\xi}(x_1), \dots, P_{n,\xi}(x_k)$ where $x_1, \dots, x_k$ is a carefully defined 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 (1.5).

It took another ten years until Ibragimov and Maslova [14, 15] successfully extended the method of Erdős–Offord to treat the Kac polynomials associated with more general distributions of ξ.

THEOREM 1.2. *For any ξ with mean zero which belongs to the domain of attraction of the normal law*,

$$\mathbf{E}N_{n,\xi} = \frac{2}{\pi}\log n + o(\log n). \tag{1.6}$$

For related results, see also [16, 17]. Few years later, Maslova [25, 26] showed that if ξ has mean zero and variance one and $\mathbf{P}(\xi = 0) = 0$, then the variance of $N_{n,\xi}$ is $(\frac{4}{\pi}(1 - \frac{2}{\pi}) + o(1))\log n$, and $N_{n,\xi}$ satisfies the central limit theorem.

So, after more than three decades of continuous research, a satisfactory answer for the Kac polynomial (the base case when all $c_i = 1$) was obtained. Apparently, the next question is what happens with *more general* sets of coefficients?

This general problem is very hard and still far from being settled. Let us recall that Kac's formula for the density function (1.2) applies for all random polynomials. However, in practice one can only evaluate this formula in the Gaussian case and some other very nice continuous distributions. On the other hand, Erdős–Offord's argument seems too delicate and relies heavily on the fact that all $c_i = 1$.

For a long time, no analogue of Theorem 1.2 was available for general sets of coefficients c_i with respect to non-Gaussian random variables ξ.

1.1. *Description of the new results for coefficients with zero means.* In this paper, we prove universality results for general random polynomials where the coefficients c_i have polynomial growth. These universality results show that, among other, the expectation of the number of real roots depend only on the mean and variance of the coefficients ξ (two moment theorems). Thus, the problem of finding the expectation of real roots reduces to the Gaussian case, which we can handle using an analytic argument (see the last paragraph of Section 3).

As the reader will see in the next section, our universality results show much more than just the expectation. They completely describe the local behavior of the roots (both complex and real). More generally, we can also control the number of intersection of the graph of the random polynomial with any deterministic curve of given degree. (The number of real roots is the number of intersections with the x-axis.)

Thanks to new and powerful tools, our method does not require an explicit expression for the deterministic coefficients c_i. As a corollary, we obtain the following extension (and refinement) of Theorem 1.2. To formulate this result (see Theorem 1.4), we first introduce a definition.

DEFINITION 1.3. We say that $h(k)$ is a generalized polynomial if there exists a finite sequence $0 < L_0 < \cdots < L_d < \infty$ such that for some $\alpha_0, \dots, \alpha_d \in \mathbb{R}$ with $\alpha_d \neq 0$ it holds that

$$h(k) = \sum_{j=0}^{d} \alpha_j \frac{L_j(L_j+1)\cdots(L_j+k-1)}{k!} \qquad \text{for every } k = 0, 1, \dots, n.$$

Here, we understand that $L \cdots (L+k-1)/k! \equiv 1$ if $k = 0$. We will say that the degree of h is $L_d - 1$ in this case. We say that h is a real generalized polynomial if the coefficients α_j's are real.

It is clear that any classical polynomial is also a generalized polynomial with the same degree: if $h(k)$ is a classical polynomial with degree d, then it could be written as a linear combination of the binomial polynomials $L_j(L_j+1)\cdots(L_j+k-1)/k!$ with $L_j = j+1$ for $j = 0, 1, \dots, d$. We also have $\alpha_d \neq 0$ because it is a nonzero multiple of the leading coefficient of h; therefore, the degree in the generalized sense of Definition 1.3 is also d. On the other hand, the class of generalized polynomials is much richer as the degree of h could be fractional.

Below, recall that our random polynomials have the form $P_n(z) = \sum_{i=0}^{n} c_i \xi_i z^i$. For a subset $S \subset \mathbb{C}$, denote by $N_P(S)$ the number of zeros of P in S.

THEOREM 1.4. *Let N_0 be a nonnegative constant. Let $\xi_0, \dots, \xi_n$ be independent (but not necessarily i.i.d.) real-valued random variables with variance 1 and $\sup_{j=0,\dots,n} \mathbf{E}|\xi_j|^{2+\epsilon} < C_0$ for some constant $C_0 > 0$ and ξ_i has mean 0 for all $i \geq N_0$. Let h be a fixed generalized polynomial with a positive leading coefficient. Assume that there are positive constants, M, m, C_1 such that the real deterministic coefficients $c_0, \dots, c_n \in \mathbb{R}$ satisfy*

$$\begin{cases} mh(k) \leq c_k^2 \leq Mh(k), & N_0 \leq k \leq n, \\ c_k^2 \leq C_1 M, & 0 \leq k < N_0. \end{cases}$$

Then with $K_h := \frac{1+\sqrt{\deg(h)+1}}{\pi}$ we have

$$\frac{m^2}{M^2}\big[K_h \log n + O(1)\big] \leq \mathbf{E}N_{P_n}(\mathbb{R}) \leq \frac{M^2}{m^2}\big[K_h \log n + O(1)\big]. \tag{1.7}$$

The implicit constants in $O(1)$ depend on ϵ, C_0, C_1, N_0, h and the ratio M/m. In particular, if $c_k^2 = h(k)$ for some real (generalized) polynomial h of degree d then

$$\mathbf{E}N_{P_n}(\mathbb{R}) = \frac{1+\sqrt{d+1}}{\pi} \log n + O(1). \tag{1.8}$$

Notice that the zeros of P_n is invariant under the scaling of c_j's, this explains why we only need dependence on the ratio M/m instead of both M and m. In the proof, we may assume $M = 1$ without loss of generality. The first few ξ_i, $i < N_0$ can have arbitrary means.

Theorem 1.4 is a corollary of our main local universality result discussed in the next section. This result (formulated in term of correlation functions) proves universality for not only the expectation, but higher moments of the number of roots (complex or real) in any small region of microscopic scale. We delay the discussion of universality to the next section and make a few comments on Theorem 1.4.

First, the error term in (1.8) is only $O(1)$, which is best possible, as showed in (1.4). Even in the well-studied case of Kac polynomials (all $c_i = 1$), this gives a improvement

$$\mathbf{E}N_{n,\xi} = \frac{2}{\pi} \log n + O(1) \tag{1.9}$$

upon the estimate $\frac{2}{\pi} \log n + o(\log n)$ from Theorem 1.2 by Ibragimov and Maslova. We believe that the method used by Erdős and Offord and also Ibragimov and Maslova cannot lead to error term better than $O(\sqrt{\log n})$. (1.9) was also proved by H. Nguyen and the last two authors in [30] by other means, but the method there does not go beyond the Kac polynomials; see also [8].

Second, there are many natural families of random polynomials which satisfy the assumptions in Theorem 1.4. Here are a few examples:

Derivatives of the Kac polynomial. The roots of the derivatives of a function have strong analytic and geometric meanings, and thus are of particular interests. For the dth derivative of the Kac polynomial (any fixed $d \geq 0$) our result implies

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

Prior to this, for derivatives of the Kac polynomial only weaker estimates [with error terms $o(\log^{1/2} n)$] are available for the Gaussian case; see the works of Das [6, 7] for $d = 1, 2$ and the extension in [35, 36] to the setting when ξ_j's are weakly correlated Gaussian random variables. For the first derivative ($d = 1$), Maslova [26] considered non-Gaussian polynomials and obtained an asymptotic bound with worse error term $o(\log n)$.

Hyperbolic polynomials. Random hyperbolic polynomials are defined by

$$c_i := \sqrt{\frac{L(L+1)\cdots(L+i-1)}{i!}},$$

for a constant $L > 0$. This class of random polynomials includes the Kac polynomials as a subcase ($L = 1$) and has became very popular recently due to the invariance of the zeros of the corresponding infinite series under hyperbolic transformations; see [13] for more discussion. By Theorem 1.4, we have

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

Logarithmic expectation. Another immediate corollary of Theorem 1.4 is that $\mathbf{E}N_{P_n}(\mathbb{R})$ grows logarithmically if the deterministic coefficients c_j have polynomial growth:

COROLLARY 1.5. *Consider ξ_i as in Theorem 1.4. Assume that there are positive constants, C_0, C_1 and some constant $\rho > -1/2$ such that the real deterministic coefficients $c_0, \dots, c_n \in \mathbb{R}$ satisfy*

$$\begin{cases} C_0 k^\rho \leq |c_k| \leq C_1 k^\rho, & N_0 \leq k \leq n, \\ c_k^2 \leq C_1, & 0 \leq k < N_0. \end{cases}$$

Then there are positive constants C_2, C_3 such that

$$C_2 \log n \leq \mathbf{E}N_{P_n}(\mathbb{R}) \leq C_3 \log n. \tag{1.10}$$

Here, C_2, C_3 depend only on C_0, C_1, ρ, N_0 and ϵ.

To deduce this result from Theorem 1.4, simply let $L = 2\rho + 1$ and notice that the binomial coefficient

$$h(k) = \frac{L(L+1)\cdots(L+k-1)}{k!}$$

is about the size of $k^{L-1} = k^{2\rho}$ for k large, therefore, the desired conclusion follows from Theorem 1.4 via comparing $|c_k|$ with $\sqrt{h(k)}$.

COROLLARY 1.6. *Consider ξ_i as in Theorem 1.4. Assume that there are positive constants, C_0, C_1 and some constant $\rho > -1/2$ such that the real deterministic coefficients $c_0, \dots, c_n \in \mathbb{R}$ satisfy*

$$\begin{cases} |c_k| = C_0 k^{\rho}\big(1+o(1)\big), & N_0 \le k \le n, \\ c_k^2 \le C_1, & 0 \le k < N_0. \end{cases}$$

Then

$$\mathbf{E} N_{P_n}(\mathbb{R}) = \frac{1+\sqrt{2\rho+1}}{\pi} \log n + o(\log n). \tag{1.11}$$

The Gaussian setting of Corollary 1.6 in the special case $c_k = k^{\rho}$, $\rho \ge 0$ was considered by [6, 7]; see also the extension in [35, 36].

To see Corollary 1.6, we need to show that given any $\delta > 0$ it holds that

$$(1-\delta)\frac{1+\sqrt{2\rho+1}}{2\pi}\log n \le \mathbf{E} N_{P_n}(\mathbb{R}) \le (1+\delta)\frac{1+\sqrt{2\rho+1}}{2\pi}\log n$$

for all n sufficiently large. Again by comparing with $\sqrt{h(k)}$ and rescaling all c_j if necessary we may assume that

$$(1+\delta)^{-1/10} h(k) \le |c_k|^2 \le (1+\delta)^{1/10} h(k)$$

for $k \le n$ sufficiently large [the threshold now depends on ρ and (polynomially) on δ]. Applying Theorem 1.4, we obtain the desired conclusion.

The reader can also notice that by Definition 1.3, our generalized polynomials always have degree greater than -1. This corresponds to the assumption that $\rho > -1/2$ in Corollaries 1.5 and 1.6. This assumption is important for our results. For example, consider the model when $c_i = i^{\rho}$ with $\rho < -1/2$ and $\mathbf{Var}\, \xi_i = 1$ for all i, then $\mathbf{Var}\, P_n(\pm 1) = \sum_{i=0}^{n} i^{2\rho}$ converges as $n \to \infty$. Intuitively, this says that the contribution of the first few terms becomes important and one may not expect to see universality around ± 1 which is where most of the real roots locate.

Number of crossings. The number of real roots is the number of intersections of the graph of $P(z)$ (over the real) with the line $y = 0$. What about an arbitrary line? (The line $y = T$ is of particular interest, as it corresponds to the important notion of level sets.) For Kac polynomials, this question was considered (see [12] for a survey) in the Gaussian case, and it was showed that the number of crossing (in expectation) is asymptotically $(\frac{2}{\pi} + o(1)) \log n$.

Theorem 1.4 allows us to prove a more precise result in much more general setting, where we can consider the number of intersection with any polynomial curve of constant degree.

COROLLARY 1.7. *Consider ξ_i as in Theorem 1.4. Assume that $c_k^2 = h(k)$ for some real (generalized) polynomial h of degree d. Let f be a deterministic polynomial of degree l and Γ be its graph over the real. Let $N_{P_n,\Gamma}(\mathbb{R})$ be the number* *of intersections of the graph of P_n (over the real) with Γ. Then*

$$\mathbf{E}N_{P_n,\Gamma}(\mathbb{R}) = \frac{1+\sqrt{d+1}}{\pi}\log n + O(1), \tag{1.12}$$

where the constant in $O(1)$ depends on ϵ, N_0, h and f.

Corollary 1.7 can be derived by applying Theorem 1.4 to the random polynomial $P_n - f$.

The Gaussian case. The strategy of the proof of Theorem 1.4 is to reduce to the Gaussian case, using universality results presented in the next section (which are the main results of this paper). Let us emphasize that even in the Gaussian setting, Theorem 1.4 (and Theorem 1.8 below) are substantially new and the method of proof is novel compared to previous works. For more details, see the last paragraph of Section 3.

1.2. *Polynomials with coefficients having nonzero means*. To conclude this section, let us mention that our method could also be used to handle polynomials with nonzero means. For instance, we have the following analogue of Theorem 1.4.

THEOREM 1.8. *Let N_0 be a positive constant and h be a deterministic classical polynomial with real coefficients. Let $\xi_0, \dots, \xi_n$ be independent real-valued random variables with variance* 1 *and* $\sup_{j=0,\dots,n} \mathbf{E}|\xi_j|^{2+\epsilon} < C_0$ *for some constant $C_0 > 0$. Assume that ξ_i has mean $\mu \neq 0$ and $c_i = h(i)$ for $i \geq N_0$ and that $|c_i| \leq C_1$ for $i < N_0$. Let $P_n(z) = \sum_{i=0}^n c_i \xi_i z^i$. Then*

$$\mathbf{E}N_{P_n}(\mathbb{R}) = \frac{1+\sqrt{2\deg(h)+1}}{2\pi}\log n + O(1). \tag{1.13}$$

The implicit constant in $O(1)$ depends on ϵ, C_0, C_1, N_0, h and μ.

The key feature of this result is that the number of real roots reduces by a factor of 2, compared to Theorem 1.4. To our best knowledge, such a result was available only for Kac polynomials. Farahmand [12] showed that when ξ is Gaussian with nonzero mean, $\mathbf{E}N_{P_n}(\mathbb{R}) = (1+o(1))\frac{1}{\pi}\log n$. Ibragimov and Maslova in [17] proved the same estimate if ξ belongs to the domain of attraction of normal law. Even for Kac polynomials, our result improves upon these as it achieves the optimal error term $O(1)$. The analogue of Corollary 1.7 holds for this model.

Similar to Theorem 1.4, Theorem 1.8 will be derived from a general universality result provided in the next section. These results can also be used to treat higher moments (such as the variance) of the number of real roots. Details will appear elsewhere.

1.3. *Outline of the paper.* In Section 2, we will present our main results regarding the local universality of the joint distribution of the zeros of $P_{n,\xi}$ when the deterministic coefficients c_j have polynomial growth. Special cases of these results will be used to reduce the proof of Theorems 1.4 and 1.8 to the Gaussian setting (see the discussion near the end of Section 2.4 for details). In Section 2.5, we will also discuss several extensions regarding universality for the zeros of random power series. Among others, we achieve local universality of hyperbolic series under very general assumptions. A sketch of our proofs for these results is presented is Section 3, followed by the detailed proofs in Sections 4, 5, 6, 7. In the rest of the paper (from Section 8 to the end), we prove the Gaussian case of Theorems 1.4 and 1.8.

It is worth mentioning that our paper is self-contained and is accessible to readers unfamiliar with the paper [38] by Tao and the third author.

2. Correlation functions and universality. Correlation functions are effective tools to study random point processes. To define correlation functions, let us first consider the complex case in which the coefficients c_i and the atom distribution ξ are not required to be real valued. In this case, the point process $\{\zeta_1, \dots, \zeta_n\}$ of zeroes of a random polynomial $P = P_n$ can be described using the (complex) *k-point correlation functions* $\rho^{(k)} = \rho_P^{(k)} : \mathbb{C}^k \to \mathbb{R}^+$, defined for any fixed natural number k by requiring that

$$\begin{aligned}(2.1)\qquad &\mathbf{E} \sum_{i_1,\dots,i_k \text{ distinct}} \varphi(\zeta_{i_1}, \dots, \zeta_{i_k}) \\ &\quad = \int_{\mathbb{C}^k} \varphi(z_1, \dots, z_k)\rho^{(k)}(z_1, \dots, z_k)\, dz_1 \cdots dz_k\end{aligned}$$

for any continuous, compactly supported, test function $\varphi : \mathbb{C}^k \to \mathbb{C}$, with the convention that $\varphi(\infty) = 0$; see, for example, [2, 13]. This definition of $\rho^{(k)}$ is clearly independent of the choice of ordering $\zeta_1, \dots, \zeta_n$ of the zeroes. Furthermore, the correlation function $\rho_P^{(k)}$ should be viewed as part of the (correlation) measure $\rho_P^{(k)}\, dz_1 \cdots dz_k$ on $\mathbb{C}^k$, and this perspective is useful in more singular settings when the correlation measure (whose existence is a consequence of the Riesz representation theorem) does not have a density on $\mathbb{C}^k$.

REMARK 2.1. When ξ has a continuous complex distribution and when the coefficients c_i are nonzero, then the zeroes are almost surely simple. In this case, if $z_1, \dots, z_k$ are distinct fixed complex numbers, then one can interpret $\rho^{(k)}(z_1, \dots, z_k)$ as the unique quantity such that the following holds: the probability that there is a zero in each of the disks $B(z_i, \varepsilon)$ for $i = 1, \dots, k$ is $(\pi\varepsilon^2)^k(\rho^{(k)}(z_1, \dots, z_k) + o(1))$ in the limit $\varepsilon \to 0$.

When the random polynomial P has real coefficients, the zeroes $\zeta_1, \dots, \zeta_n$ are symmetric with respect to the real axis, and one expects several of the zeroes to lie on this axis. Because of this possibility, the situation is more complicated. It is no longer natural to work with the complex k-point correlation functions $\rho_P^{(k)}$, as they are likely to become singular on the real axis. Instead, we divide the complex plane $\mathbb{C}$ into three pieces $\mathbb{C} = \mathbb{R} \cup \mathbb{C}_+ \cup \mathbb{C}_-$, with $\mathbb{C}_+ := \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ being the upper half-plane and $\mathbb{C}_- := \{z \in \mathbb{C} : \mathrm{Im}(z) < 0\}$ being the lower half-plane. By the aforementioned symmetry, we may restrict our attention to the zeroes in $\mathbb{R}$ and $\mathbb{C}_+$ only. For any natural numbers $k, l \geq 0$, we define the *mixed* (k,l)*-correlation function* $\rho^{(k,l)} = \rho_P^{(k,l)} : \mathbb{R}^k \times (\mathbb{C}_+ \cup \mathbb{C}_-)^l \to \mathbb{R}^+$ of a random polynomial P to be the function defined by the formula

$$\mathbf{E} \sum_{i_1,\dots,i_k \text{ distinct}} \sum_{j_1,\dots,j_l \text{ distinct}} \varphi(\zeta_{i_1,\mathbb{R}}, \dots, \zeta_{i_k,\mathbb{R}}, \zeta_{j_1,\mathbb{C}_+}, \dots, \zeta_{j_l,\mathbb{C}_+})$$

$$= \int_{\mathbb{R}^k} \int_{\mathbb{C}_+^l} \varphi(x,z) \rho_P^{(k,l)}(x,z)\, dz\, dx \qquad \big(x \in \mathbb{R}^k, z \in \mathbb{C}_+^l\big)$$

for any continuous compactly supported test function $\varphi : \mathbb{R}^k \times \mathbb{C}^l \to \mathbb{C}$ (note that we do not require φ to vanish at the boundary of $\mathbb{C}_+^l$), $\zeta_{i,\mathbb{R}}$ runs over an arbitrary enumeration of the real zeroes of P_n, and $\zeta_{j,\mathbb{C}_+}$ runs over an arbitrary enumeration of the zeroes of P_n in $\mathbb{C}_+$. This defines $\rho^{(k,l)}$ (in the sense of distributions, at least) for $x_1, \dots, x_k \in \mathbb{R}$ and $z_1, \dots, z_l \in \mathbb{C}_+$; we then extend $\rho^{(k,l)}(x_1, \dots, x_k, z_1, \dots, z_l)$ to all other values of $x_1, \dots, x_k \in \mathbb{R}$ and $z_1, \dots, z_l \in \mathbb{C}_+ \cup \mathbb{C}_-$ by requiring that $\rho^{(k,l)}$ is symmetric with respect to conjugation of any or all of the $z_1, \dots, z_l$ parameters. Again, we permit $\rho^{(k,l)}$ to be a measure[3] instead of a function when the random polynomial P_n has a discrete distribution.

In the case $l = 0$, the correlation functions $\rho^{(k,0)}$ for $k \geq 1$ provide (in principle, at least) all the essential information about the distribution of the real zeroes. For instance,

$$\mathbf{E} N_P(\mathbb{R}) = \int_{\mathbb{R}} \rho^{(1,0)}(x)\, dx \tag{2.2}$$

and similarly,

$$\begin{aligned} \mathbf{Var}\, N_P(\mathbb{R}) = &\int_{\mathbb{R}} \int_{\mathbb{R}} \rho^{(2,0)}(x,y) - \rho^{(1,0)}(x) \rho^{(1,0)}(y)\, dx\, dy \\ &+ \int_{\mathbb{R}} \rho^{(1,0)}(x)\, dx. \end{aligned} \tag{2.3}$$

We refer the reader to [2, 13] for a thorough discussion of correlation functions.

[3]As in the complex case, we allow the real zeros $\zeta_{i_1,\mathbb{R}}, \dots, \zeta_{i_k,\mathbb{R}}$ or the complex zeroes $\zeta_{j_1,\mathbb{C}_+}, \dots, \zeta_{j_l,\mathbb{C}_+}$ to have multiplicity; it is only the indices $i_1, \dots, i_k, j_1, \dots, j_l$ that are required to be distinct. In particular, in the discrete case it is possible for $\rho^{(0,2)}(z_1, z_2)$ (say) to have nonzero mass on the diagonal $z_1 = z_2$ or the conjugate diagonal $z_1 = \overline{z_2}$, if P has a repeated complex eigenvalue with positive probability.

2.1. *Universality.* The correlation functions give us a lot of information at finer scales. Given the story of real roots in the previous section [which corresponds to the special case (2.2)], it is natural to expect that their computation is extremely hard.

The situation is roughly as follows. We have an explicit formula (Kac–Rice formula) to compute correlation functions. This formula is a generalization of Kac's formula in the previous section and involves joint distributions. In principle, one can evaluate it in the Gaussian case (as Kac did). But technically, for various sets of coefficients c_i, this is already a significant challenge.

In [29], Nazarov and Sodin considered the series $f(z) = \sum_{j=0}^{\infty} \frac{1}{\sqrt{j!}}\xi_j z^j$ where ξ_j are i.i.d. normalized complex Gaussian and used the Kac–Rice formula to prove repulsion properties of its complex zeros, more specifically they proved that the k-correlation function is locally comparable to the square modulus of the complex Vandermonde product:

$$C^{-1}\prod_{i<j}|z_i - z_j|^2 \le \rho_f^{(k)}(z_1, \dots, z_k) \le C\prod_{i<j}|z_i - z_j|^2.$$

The method of [29] extends to more general settings. In [29], the authors proved the same type of estimates for the complex k-point correlation function of the so-called $2k$-nondegenerate Gaussian analytic functions, which include (among others) $P(z) = \sum_{j=0}^{\infty} c_j\xi_j z^j$ with $c_0, \dots, c_{2k-1} \neq 0$ such that $\sum_j |c_j|^2|z|^{2j}$ converges in the domain where estimates for $\rho_f^{(k)}$ are needed (see [29] for technical details). These certainly include random polynomials of finite degrees (at least $2k-1$) whose first $2k$ coefficients are nonzero; however, the implicit constants C in the estimates depend also on f (and k and the domain), and thus could be a very large function of the degree.

Similar to the Kac formula, a direct evaluation of the Kac–Rice formula is not feasible when ξ is a general non-Gaussian random variable. On the other hand, it has been conjectured that the value of the formula, at least in the asymptotic sense, should not depend on the fine details of the atom variable ξ. This is commonly referred to in the literature as the *universality phenomenon*.

Bleher and Di proved universality for elliptic polynomials in which the atom distribution ξ was real-valued and sufficiently smooth and rapidly decaying (see [4], Theorem 7.2, for the precise technical conditions and statement). With these hypotheses, they showed that the pointwise limit of the normalized correlation function $n^{-k/2}\rho^{(k,0)}(a + \frac{x_1}{\sqrt{n}}, \dots, a + \frac{x_k}{\sqrt{n}})$ for any fixed $k, a, x_1, \dots, x_k$ (with $a \neq 0$) as $n \to \infty$ was independent of the choice of ξ (with an explicit formula for the limiting distribution). Their method is based on the Kac–Rice formula.

In a recent paper [38], Tao and the third author introduced a new method to prove universality, which we will refer to as "universality by sampling" (see Section 3). This method makes no distinction between continuous and discrete random variables and the authors used it to derive universality for flat, elliptic and Kac polynomials in certain domains.

DEFINITION 2.2. Two complex random variables ξ and ξ' are said to *match moments to order m* if

$$\mathbf{E}\mathrm{Re}(\xi)^a \,\mathrm{Im}(\xi)^b = \mathbf{E}\mathrm{Re}(\xi')^a \,\mathrm{Im}(\xi')^b$$

for all natural numbers $a, b \geq 0$ with $a + b \leq m$.

2.2. *Coefficients with polynomial growth.* We consider

$$P_n(z) = \sum_{i=0}^{n} c_i \xi_i z^i, \qquad z \in \mathbb{C}, \tag{2.4}$$

random polynomials with the following condition.

CONDITION 1. 1. ξ_i's are independent (real or complex) random variables with unit variance and bounded $2 + \epsilon$ moment, namely $\mathbf{E}|\xi_i|^{2+\epsilon} \leq \tau_2$ for an arbitrarily small positive constant ϵ.

2. c_i's are deterministic complex numbers with

$$\tau_1 i^\rho \leq |c_i| \leq \tau_2 i^\rho \qquad \text{for all } i \geq N_0, \text{ and } |c_i| \leq \tau_2 \text{ for all } 0 \leq i < N_0, \tag{2.5}$$

where N_0, τ_1, τ_2, ϵ are positive constants and $\rho > -1/2$.

Notice that we do not require the ξ_i to be identically distributed. They are also allowed to have different means. However, by Hölder's inequality, our condition on the uniform boundedness of $2 + \epsilon$ moments implies that the means should be bounded $\mathbf{E}|\xi_i| \leq \tau_2^{1/(2+\epsilon)}$ for all i.

An essential point here is that we do not need to know the values of the coefficients c_i precisely, only their growth. We do not know of any result which is applicable at this level of generality.

In the next two subsections, we state our universality theorems for complex and mixed correlation functions.

2.3. *Complex local universality for polynomials.* For a polynomial $P = P_n$ of the form (2.4), let $(\zeta_i^P)_{i=1}^n$ be the zeros of P. We use the convention that if P_n vanishes identically then it has a zero of order n at ∞, and similarly, if P_n has degree $m < n$ then it has a zero of order $n - m$ at ∞.

Let δ be a small positive number. Define

$$I(\delta) = \begin{cases} [1 - 2\delta, 1 - \delta] & \text{if } \dfrac{1}{10n} \leq \delta < 1, \\ [1 - 1/n, 1 + 1/n] & \text{if } 0 < \delta < \dfrac{1}{10n}, \end{cases}$$

and

$$J(\delta) = \left[\frac{1}{1 - \delta}, \frac{1}{1 - 2\delta}\right] \qquad \text{if } \frac{1}{10n} \leq \delta < 1. \tag{2.6}$$

Note that

$$\Big(\bigcup_{\frac{1}{20n}\le\delta\le\frac{1}{C}} I(\delta)\Big)\cup\Big(\bigcup_{\frac{1}{10n}\le\delta\le\frac{1}{C}} J(\delta)\Big)=\Big[1-\frac{2}{C},\frac{C}{C-2}\Big]\supset\Big[1-\frac{1}{C},1+\frac{1}{C}\Big].$$

Our goal is to prove universality in the annulus $A(0, 1-\frac{1}{C}, 1+\frac{1}{C})$ for some large constant C and we shall break it into annuli with radii given by $I(\delta)$ and $J(\delta)$. When proving universality on the annulus $\{z : |z| \in I(\delta)\}$, for convenience of notation we will consider the following rescaled version:

$$\check{P}(\check{z}) = P(z) \qquad \text{where } \check{z} = \frac{z}{10^{-3}\delta}. \tag{2.7}$$

The term "local universality" can be thought of as universality on balls that contain $\Theta(1)$ zeros on average. It is more or less proven throughout the paper that for such z as above, there are an average of $\Theta(1)$ zeros in the ball $B(z, 10^{-3}\delta)$. The rescaled factor $10^{-3}\delta$ in (2.7) plays the simple role of making this ball have the unit radius. The factor 10^{-3} is artificial and can be replaced by any sufficient small constant that allows the ball to grow under various approximation steps in our proofs while keeping distance $\Theta(\delta)$ away from the unit circle. Observe that by the change of variables formula, we have

$$\rho^{(k)}_{\check{P}}(w_1,\dots,w_k) = \big(10^{-3}\delta\big)^{2k}\rho^{(k)}_{P}\big(10^{-3}\delta w_1,\dots,10^{-3}\delta w_k\big).$$

When working with the annulus $\{z : |z| \in J(\delta)\}$, we first consider $Q(z) = \frac{z^n}{c_n}P(\frac{1}{z})$ to transform the domain $|z| \ge 1$ into $|z| \le 1$, and in particular, $J(\delta)$ into $I(\delta)$. Note that $Q = \sum_{i=0}^{n} \frac{d_i}{d_0}\xi_{n-i} z^i$ where $d_i = c_{n-i}$. For notational convenience, sometimes we also think about Q as $Q = \sum_{i=0}^{n} \frac{d_i}{d_0}\xi_i z^i$. And then, we use the same rescaling:

$$\check{Q}(\check{z}) = Q(z) \qquad \text{where } \check{z} = \frac{z}{10^{-3}\delta}.$$

Let $\rho^{(k)}_{\check{P}}$ and $\rho^{(k,l)}_{\check{P}}$ be the corresponding correlation functions of $\check{P}$. Note that they depend on δ because the rescaling factor does.

Here, for a function $F : \mathbb{C}^m \to \mathbb{C}$, we think of it as a function from $\mathbb{R}^{2m} \to \mathbb{C}$ and denote by $|\nabla^a F(x)|$ the Euclidean norm of $\nabla^a F(x)$:

$$\big|\nabla^a F(x)\big| = \Big(\sum_{1\le i_1,\dots,i_a\le 2m}\Big|\frac{\partial^a F}{\partial x_{i_1}\cdots\partial x_{i_a}}(x)\Big|^2\Big)^{1/2}.$$

THEOREM 2.3. *Let $k \ge 1$ be an integer constant. Let $P_n = \sum_{i=0}^{n} c_i\xi_i z^i$ and $\widetilde{P}_n = \sum_{i=0}^{n} c_i\tilde{\xi}_i z^i$ be two random polynomials satisfying Condition 1. Assume that ξ_i and $\tilde{\xi}_i$ match moments to second order for all $N_0 \le i \le n$ where N_0 is the constant in Condition 1.*

Let $\Delta\rho_P^{(k)} = \rho_{\check{P}}^{(k)} - \rho_{\check{\tilde{P}}}^{(k)}$ *and* $\Delta\rho_Q^{(k)} = \rho_{\check{Q}}^{(k)} - \rho_{\check{\tilde{Q}}}^{(k)}$, *differences of correlation functions.*

Then there exist constants C, C', c *depending only on* k *and the constants in Condition* 1 *such that for every* $\frac{1}{20n} \le \delta \le \frac{1}{C}$ *and complex numbers* $z_1, \dots, z_k$ *with* $|z_j| \in I(\delta)$ *for all* $0 \le j \le k$, *and for every smooth function* $G : \mathbb{C}^k \to \mathbb{C}$ *supported on* $B(0, 10^{-3})^k$ *with* $|\nabla^a G(z)| \le 1$ *for all* $0 \le a \le 2k+4$ *and* $z \in \mathbb{C}^k$, *we have*

$$\left| \int_{\mathbb{C}^k} G(w_1, \dots, w_k) \Delta\rho_P^{(k)}(\check{z}_1 + w_1, \dots, \check{z}_k + w_k)\, dw_1 \cdots dw_k \right| \le C' \delta^c. \tag{2.8}$$

Furthermore, if $\frac{1}{10n} \le \delta \le \frac{1}{C}$,

$$\left| \int_{\mathbb{C}^k} G(w_1, \dots, w_k) \Delta\rho_Q^{(k)}(\check{z}_1 + w_1, \dots, \check{z}_k + w_k)\, dw_1 \cdots dw_k \right| \le C' \delta^c. \tag{2.9}$$

2.4. *Real local universality.* For real universality, we require the following additional condition on ξ_i's and c_i's.

CONDITION 2. 1. The random variables ξ_i's and the coefficients c_i's are real.
2. One of the following holds:

(a) $\mathbf{E}\xi_i = 0$ for all $i \ge N_0$,
(b) $\mathbf{E}\xi_i = \mu$ for all $i \ge N_0$, where μ is any constant, and there exists a classical polynomial $\mathfrak{P}$ (independent of n) with degree $\rho \in \mathbb{N}$ such that $c_i = \mathfrak{P}(i)$ for all $i \ge N_0$.[4]

Notice that when Condition 2(2b) is satisfied, by replacing c_i by $-c_i$ if needed, we can also assume that $c_i = \mathfrak{P}(i) > 0$ for all i larger than some constant because the (fixed) polynomial $\mathfrak{P}(x)$ keeps the same sign when x is sufficiently large.

THEOREM 2.4. *Let* $k, l \ge 0$ *be integer constants with* $k + l \ge 1$. *Let* $P_n = \sum_{i=0}^n c_i \xi_i z^i$ *and* $\widetilde{P}_n = \sum_{i=0}^n c_i \tilde{\xi}_i z^i$ *be two random polynomials satisfying Conditions* 1 *and* 2. *Assume that* ξ_i *and* $\tilde{\xi}_i$ *match moments to second order for all* $N_0 \le i \le n$ *where* N_0 *is the constant in Condition* 1.

Let $\Delta\rho_P^{(k)} = \rho_{\check{P}}^{(k)} - \rho_{\check{\tilde{P}}}^{(k)}$ *and* $\Delta\rho_Q^{(k)} = \rho_{\check{Q}}^{(k)} - \rho_{\check{\tilde{Q}}}^{(k)}$, *differences of correlation functions.*

Then there exist constants C, c *depending only on* k, l *and the constants and the polynomial* $\mathfrak{P}$ *in Conditions* 1 *and* 2 *such that for every* $\frac{1}{20n} \le \delta \le \frac{1}{C}$, *real numbers* $x_1, \dots, x_k$, *and complex numbers* $z_1, \dots, z_l$ *such that* $|x_i|, |z_j| \in I(\delta)$ *for all* $i = 1, \dots, k$, $j = 1, \dots, l$, *and for every smooth function* $G : \mathbb{R}^k \times \mathbb{C}^l \to \mathbb{C}$

[4]For instance, P is Kac polynomial or its derivatives.

supported on $[-10^{-3}, 10^{-3}]^k \times B(0, 10^{-3})^l$ *such that* $|\nabla^a G(z)| \le 1$ *for all* $0 \le a \le 2(k+l)+4$ *and* $z \in \mathbb{R}^k \times \mathbb{C}^l$, *we have*

$$\begin{aligned}\Big|\int_{\mathbb{R}^k}\int_{\mathbb{C}^l} & G(y_1,\dots,y_k,w_1,\dots,w_l) \\ & \times \Delta\rho_P^{(k,l)}(\check{x}_1+y_1,\dots,\check{x}_k+y_k, \\ & \check{z}_1+w_1,\dots,\check{z}_l+w_l)\,dy_1\cdots dy_k\,dw_1\cdots dw_l\Big| \\ \le{} & C\delta^c. \end{aligned} \tag{2.10}$$

Furthermore, if $\frac{1}{10n} \le \delta \le \frac{1}{C}$, *we have the same inequality* (2.10) *with* Q *in place of* P.

Now, to derive Theorems 1.4 and 1.8 from Theorem 2.4, it suffices to show that the number of real roots in the Gaussian case satisfies the claimed bounds and that the expectation of real roots (in the general case) outside the universality annulus is bounded. More specifically, we will show the following.

LEMMA 2.5. *Under the conditions of Theorem* 2.4, *for each constant* $C > 0$, *there exists a constant* $M(C)$ *such that*

$$\mathbf{E}N_{P_n}\left(\mathbb{R}\setminus A\left(0, 1-\frac{1}{C}, 1+\frac{1}{C}\right)\right) \le M(C),$$

for every $n \ge 1$.

Together with Theorem 2.4, Lemma 2.5 gives the following.

COROLLARY 2.6. *Under conditions of Theorem* 2.4, *there exists a constant* C *such that for every* $n \ge 1$, *one has*

$$\left|\mathbf{E}N_{P_n}(\mathbb{R}) - \mathbf{E}N_{\tilde{P}_n}(\mathbb{R})\right| \le C.$$

REMARK 2.7. To get an intuition for Lemma 2.5, let i be the smallest index for which $c_{i_0} \ne 0$. Assume that $c_{i_0} = \Omega(1)$, $\mathbf{E}\log|\xi_{i_0}| = O(1)$, and $\mathbf{E}\log|\xi_n| = O(1)$. Under condition (2.5), $i_0 = O(1)$. Then by Jensen's inequality for the function $P(z)/z^{i_0}$ and concavity of the function log, one has the easy bound

$$\begin{aligned}\mathbf{E}N_P\big(B(0,1-1/C)\big) &\le i_0 + \mathbf{E}\frac{\log\frac{M}{|c_{i_0}\xi_{i_0}|}}{\log\frac{1-1/2C}{1-1/C}} \\ &= i_0 + O_C(1) + O_C(\mathbf{E}\log M) \\ &\le O_C(1) + O_C(\log\mathbf{E}M) \\ &= O_C(1) + O_C\left(\log\sum_{i=0}^{\infty} i^\rho(1-1/2C)^i\right) = O_C(1),\end{aligned}$$

where $M = \max_{|z|\le 1-1/2C} |P(z)/z^{i_0}|$. Similarly, $\mathbf{E}N_Q(B(0, 1-1/C)) = O_C(1)$. And hence, $\mathbf{E}N_P(\mathbb{C}\setminus A(0, 1-1/C, 1+1/C)) = O_C(1)$. In other words, for a large class of polynomials of the form (2.4), one expects to see only a few zeros outside the annulus of universality.

By Corollary 2.6, to verify Theorem 1.4 and Theorem 1.8 it suffices to consider the Gaussian case.

THEOREM 2.8. *The statement of Theorem 1.4 holds for ξ_i being standard Gaussian for all $i = 0, \dots, n$.*

THEOREM 2.9. *The statement of Theorem 1.8 holds for ξ_i being Gaussian with mean μ and variance 1 for all $i = 0, \dots, n$.*

We are going to prove these theorems in Section 8 and Section 11. The evaluation of Kac's formula under the general setting of Theorem 1.4 is fairly involved, and as mentioned in the discussion leading to Corollary 1.5, it is somewhat surprising that the growth of the coefficients alone already determines the number of real roots.

2.5. *Local universality for series.* Our method could also be used to extend the previous results to random series. Let us first extend Theorem 2.3.

We consider a random series P_{PS} of the form

$$P_{\mathrm{PS}}(z) = \sum_{i=0}^{\infty} c_i \xi_i z^i, \qquad z \in \mathbf{D}, \tag{2.11}$$

where $\mathbf{D}$ is the open unit disk in the complex plane, and the c_i's and ξ_i's satisfy Condition 1.

THEOREM 2.10. *Let $k \ge 1$ be an integer constant. Let $P_{\mathrm{PS}} = \sum_{i=0}^{\infty} c_i \xi_i z^i$ and $\widetilde{P}_{\mathrm{PS}} = \sum_{i=0}^{\infty} c_i \tilde{\xi}_i z^i$ be two random power series satisfying Condition 1 (with n being replaced by ∞). Assume that ξ_i and $\tilde{\xi}_i$ match moments to second order for all $i \ge N_0$ where N_0 is the constant in Condition 1.*

Let $\Delta\rho^{(k)} = \rho^{(k)}_{\check{P}_{\mathrm{PS}}} - \rho^{(k)}_{\check{\tilde{P}}_{\mathrm{PS}}}$, the difference of the correlation functions.

Then there exist constants C, c depending only on k and the constants in Condition 1 such that for every $0 < \delta \le \frac{1}{C}$ and complex numbers $z_1, \dots, z_k$ with $|z_j| \in [1-2\delta, 1-\delta]$ for all $0 \le j \le k$, and for every smooth function $G : \mathbb{C}^k \to \mathbb{C}$ supported on $B(0, 10^{-3})^k$ with $|\nabla^a G(z)| \le 1$, $\forall 0 \le a \le 2k+4$ and $z \in \mathbb{C}^k$, we have

$$\left| \int_{\mathbb{C}^k} G(w_1, \dots, w_k) \Delta\rho^{(k)}(\check{z}_1 + w_1, \dots, \check{z}_k + w_k)\, dw_1 \cdots dw_k \right| \le C\delta^c. \tag{2.12}$$

Notice that when all ξ_i are (complex) standard Gaussian, the distribution of the zeroes is invariant with respect to rotation. As a corollary of Theorem 2.10, this invariance is preserved (in the asymptotic sense) if ξ_i matches the moments of standard Gaussian up to second order.

COROLLARY 2.11. *Let $k \geq 1$ be an integer constant. Let P_{PS} be the random series of the form (2.11) satisfying Condition 1. Assume furthermore that* $\mathbf{E}(\mathrm{Re}(\xi_i)) = \mathbf{E}(\mathrm{Im}(\xi_i)) = \mathrm{Cov}(\mathrm{Re}(\xi_i), \mathrm{Im}(\xi_i)) = 0$ *and* $\mathbf{Var}(\mathrm{Re}(\xi_i)) = \mathbf{Var}(\mathrm{Im}(\xi_i)) = 1/2$ *for all $i \geq N_0$.*

Then there exist constants C, c such that for every $0 < \delta \leq \frac{1}{C}$ and complex numbers $z_1, \dots, z_k$ with $|z_j| \in [1-2\delta, 1-\delta]$ for all $0 \leq j \leq k$ and $0 \leq \theta < 2\pi$, and for every smooth function $G : \mathbb{C}^k \to \mathbb{C}$ supported on $B(0, 10^{-3})^k$ with $|\nabla^a G(z)| \leq 1$, $\forall 0 \leq a \leq 2k+4$ and $z \in \mathbb{C}^k$, we have

$$\begin{aligned}
&\bigg| \int_{\mathbb{C}^k} G(w_1, \dots, w_k) \rho^{(k)}_{\check{P}_{\mathrm{PS}}}(\check{z}_1 + w_1, \dots, \check{z}_k + w_k)\, dw_1 \cdots dw_k \\
&\qquad - \int_{\mathbb{C}^k} H(w_1, \dots, w_k) \\
&\qquad \times \rho^{(k)}_{\check{P}_{\mathrm{PS}}}\big(e^{\sqrt{-1}\theta} \check{z}_1 + w_1, \dots, e^{\sqrt{-1}\theta} \check{z}_k + w_k\big)\, dw_1 \cdots dw_k \bigg| \\
&\quad \leq C\delta^c,
\end{aligned}$$

where $H(w_1, \dots, w_k) = G(e^{-\sqrt{-1}\theta} w_1, \dots, e^{-\sqrt{-1}\theta} w_k)$.

In case that P_{PS} is hyperbolic and the ξ_i are complex Gaussian, the distribution of the zeros of P_{PS} is invariant under hyperbolic transformations of the disk $\mathbf{D}$ (see [13]). A hyperbolic transformation on $\mathbf{D}$ is a transformation of the form

$$\phi(z) = \frac{az + b}{\bar{b}z + \bar{a}},$$

where $a, b \in \mathbb{C}$ and $|a|^2 - |b|^2 = 1$. A holomorphic function on $\mathbf{D}$ is bijective if and only if it is a hyperbolic transformation (see, for instance, [34], Theorems 12.4, 12.6).

As another immediate corollary of Theorem 2.10, this invariance is preserved (in the asymptotic sense) again if ξ_i matches the moments of standard Gaussian up to order 2 and if the hyperbolic transformation preserves our universality domain.

COROLLARY 2.12. *Let $k \geq 1$ be an integer constant. Let P be the random* hyperbolic *series of the form (2.11) satisfying Condition 1. Assume furthermore that* $\mathbf{E}(\mathrm{Re}(\xi_i)) = \mathbf{E}(\mathrm{Im}(\xi_i)) = \mathrm{Cov}(\mathrm{Re}(\xi_i), \mathrm{Im}(\xi_i)) = 0$ *and* $\mathbf{Var}(\mathrm{Re}(\xi_i)) = \mathbf{Var}(\mathrm{Im}(\xi_i)) = 1/2$ *for all $i \geq N_0$.*

Then there exist constants C, c such that the following holds true. Let $0 < \delta_0 \leq \frac{1}{C}$ and complex numbers $z_1, \dots, z_k$ with $|z_j| \in [1 - 2\delta_0, 1 - \delta_0]$ for all $0 \leq j \leq k$

and $0 \le \theta < 2\pi$. *Let* ϕ *be a hyperbolic transformation that maps* z_j *to* t_j *with* $|t_j| \in [1-2\delta_1, 1-\delta_1]$ *for all* j *and for some* $0 < \delta_1 \le \frac{1}{C}$. *Then for every smooth function* $G : \mathbb{C}^k \to \mathbb{C}$ *supported on* $B(0, 10^{-4})^k$ *with* $|\nabla^a G(z)| \le 1$, $\forall 0 \le a \le 2k+4$ *and* $z \in \mathbb{C}^k$, *we have*

$$\left| \int_{\mathbb{C}^k} G(w) \left(\frac{\delta_0}{10^3} \right)^{2k} \rho^{(k)}_{P_{\mathrm{PS}}} (z + 10^{-3}\delta_0 w) \, dw_1 \cdots dw_k \right. \\ \left. - \int_{\mathbb{C}^k} H(w) (10^{-3}\delta_1)^{2k} \rho^{(k)}_{P_{\mathrm{PS}}} (t + 10^{-3}\delta_1 w) \, dw_1 \cdots dw_k \right| \le C \max\{\delta_0, \delta_1\}^c,$$

where

$$H(w_1, \ldots) = G\left(\frac{1}{10^{-3}\delta_0} (\phi^{-1}(t_1 + 10^{-3}\delta_1 w_1) - z_1), \ldots \right).$$

Similar to the complex case, real universality also follows from our arguments for polynomials.

THEOREM 2.13. *Let* $k, l \ge 0$ *be integer constants with* $k + l \ge 1$. *Let* $P_{\mathrm{PS}} = \sum_{i=0}^{\infty} c_i \xi_i z^i$ *and* $\tilde{P}_{\mathrm{PS}} = \sum_{i=0}^{\infty} c_i \tilde{\xi}_i z^i$ *be two random power series satisfying Conditions* 1 *and* 2 (*with* n *being replaced by* ∞). *Assume that* ξ_i *and* $\tilde{\xi}_i$ *match moments to second order for all* $i \ge N_0$ *where* N_0 *is the constant in Conditions* 1 *and* 2.

Let $\Delta\rho^{(k,l)} = \rho^{(k,l)}_{\check{P}_{\mathrm{PS}}} - \rho^{(k,l)}_{\check{\tilde{P}}_{\mathrm{PS}}}$, *the difference of the correlation functions.*

Then there exist constants C, c *depending only on* k, l *and the constants and the polynomial* $\mathfrak{P}$ *in Conditions* 1 *and* 2 *such that for every* $0 < \delta \le \frac{1}{C}$, *real numbers* $x_1, \ldots, x_k$, *and complex numbers* $z_1, \ldots, z_l$ *such that* $|x_i|, |z_j| \in [1 - 2\delta, 1 - \delta]$ *for all* $i = 1, \ldots, k$, $j = 1, \ldots, l$, *and for every smooth function* $G : \mathbb{R}^k \times \mathbb{C}^l \to \mathbb{C}$ *supported on* $[-10^{-3}, 10^{-3}]^k \times B(0, 10^{-3})^l$ *such that* $|\nabla^a G(z)| \le 1$, $\forall 0 \le a \le 2(k+l)+4$ *and* $z \in \mathbb{R}^k \times \mathbb{C}^l$, *we have*

$$\left| \int_{\mathbb{R}^k} \int_{\mathbb{C}^l} G(y_1, \ldots, y_k, w_1, \ldots, w_l) \right. \\ \times \Delta\rho^{(k,l)} (\check{x}_1 + y_1, \ldots, \check{x}_k + y_k, \\ \left. \check{z}_1 + w_1, \ldots, \check{z}_l + w_l) \, dy_1 \cdots dy_k \, dw_1 \cdots dw_l \right| \\ \le C\delta^c. \tag{2.13}$$

We will prove these results in Section 7.

3. Sketch of the proof and the main technical ideas. To start, we make use of the "universality by sampling" method from [38], which is based on the Lindeberg swapping technique. To give the reader a quick introduction on this method, let us discuss the simplest correlation function $\rho^{(0,1)}$, which is the density function of the complex roots. Consider two polynomials $P_{n,\xi}$ and $P_{n,\tilde{\xi}}$ and a (nice) test function $G(x)$. We would like to show

$$\int_{\mathbb{C}} G(x)\rho^{(0,1)}_{P_{n,\xi}}(x)\,dx = \int_{\mathbb{C}} G(x)\rho^{(0,1)}_{P_{n,\tilde{\xi}}}(x)\,dx + o(1).$$

Recall that by definition

$$\int_{\mathbb{C}} G(x)\rho^{(0,1)}_{P_{n,\xi}}(x)\,dx = \sum_{i=1}^{n} \mathbf{E}_{\xi} G(\zeta_i),$$

$$\int_{\mathbb{C}} G(x)\rho^{(0,1)}_{P_{n,\tilde{\xi}}}(x)\,dx = \sum_{i=1}^{n} \mathbf{E}_{\tilde{\xi}} G(\tilde{\zeta}_i),$$

where ζ_i ($\tilde{\zeta}_i$) are the roots of $P_{n,\xi}$ ($P_{n,\tilde{\xi}}$).

We are going to prove universality of the right-hand side, namely

$$\sum_{i=1}^{n} \mathbf{E}_{\xi} G(\zeta_i) = \sum_{i=1}^{n} \mathbf{E}_{\tilde{\xi}} G(\tilde{\zeta}_i) + o(1).$$

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

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

where Δ is the Laplacian. By change of variables, this implies that for all i,

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

which, in turn, yields

$$\begin{aligned}\sum_i \mathbf{E}_{\xi} G(\zeta_i) &= -\frac{1}{2\pi} \mathbf{E}_{\xi} \int_{\mathbb{C}} \log\left|\prod_{i=1}^{n}(z-\zeta_i)\right| \Delta G(z)\,dz \\ &= -\frac{1}{2\pi} \mathbf{E}_{\xi} \int_{\mathbb{C}} \log\bigl|P_{n,\xi}(z)\bigr| \Delta G(z)\,dz.\end{aligned}$$

[The leading coefficient of $P_{n,\xi}$ does not matter here, as $\int_{\mathbb{C}} \Delta G(z)\,dz = 0$.] We estimate the integration $\int_{\mathbb{C}} \log|P_{n,\xi}(z)|\Delta G(z)\,dz$ by *sampling*. The intuition is that if S is the average of (say) N numbers $S := \frac{a_1+\cdots+a_N}{N}$ where N is large integer, then (hopefully) we can estimate S accurately by a much shorter random partial sum $S' = \frac{a_{i_1}+\cdots+a_{i_m}}{m}$, where the indices $i_1, \dots, t_m$ are chosen randomly from the index set $\{1, \dots, N\}$, with m being a parameter much smaller than N. Thinking of $a_1, \dots, a_N$ as terms in the Riemann sum approximation of $\int_{\mathbb{C}} \log|P_{n,\xi}(z)|\Delta G(z)\,dz$, we want to approximate this integral by

$$\frac{1}{m}\bigl(H_{\xi}(z_1) + \cdots + H_{\xi}(z_m)\bigr),$$

where $H_\xi(z) := C\log|P_{n,\xi}(z)|\Delta G(z)$ with C being a normalization constant, $z_1,\dots,z_m$ are random sample points, and m is a properly chosen parameter which tends to infinity slowly with n. (The magnitude of m determines the quality of the approximation.)

Now assume, for a moment that $\frac{1}{m}(H_\xi(z_1)+\cdots+H_\xi(z_m))$ is indeed a good approximation of $\int_{\mathbb{C}}\log|P_{n,\xi}(z)|\Delta G(z)\,dz$, and similarly, $\frac{1}{m}(H_{\tilde\xi}(z_1)+\cdots+H_{\tilde\xi}(z_m))$ is a good approximation of $\int_{\mathbb{C}}\log|P_{n,\xi}(z)|\Delta G(z)\,dz$, with overwhelming probability. In this case, the problem reduces to showing

$$\mathbf{E}_\xi\frac{1}{m}\big(H_\xi(z_1)+\cdots+H_\xi(z_m)\big)=\mathbf{E}_{\tilde\xi}\frac{1}{m}\big(H_{\tilde\xi}(z_1)+\cdots+H_{\tilde\xi}(z_m)\big)+o(1).$$

We can apply the Lindeberg swapping method to prove this estimate. In fact, we can use this method to show that the joint distribution of m variables $H_\xi(z_1),\dots,H_\xi(z_m)$ and that of $H_{\tilde\xi}(z_1),\dots,H_{\tilde\xi}(z_m)$ are approximately the same. This can be done by defining $Z:=(H_\xi(z_1),\dots,H_\xi(z_m))$ and showing

$$\mathbf{E}_\xi F(Z)=\mathbf{E}_{\tilde\xi}F(\widetilde{Z})+o(1) \tag{3.1}$$

for any nice test function F.

An application of the Lindeberg method often requires estimates on the derivatives of the function in question, and a decisive advantage here is that the function H is explicit, and it is not too hard to bound its derivatives. Generalizing the whole scheme to the general case of $\rho^{k,l}$ requires several additional technical steps, but the spirit of the method remains the same.

The critical point of this scheme is to show that the random sum indeed approximates the integral. In order to do so, we need to bound from above the second moment

$$\int_{\mathbb{C}}\big|\log\big|P_{n,\xi}(z)\big|\Delta G(z)\big|^2\,dz=\int_D\big|\log\big|P_{n,\xi}(z)\big|\Delta G(z)\big|^2\,dz,$$

where D is the support of G; see Lemma 4.8.

Our strategy has two steps. We first define a *good* event $\mathcal{T}$ (which holds with high probability) in the space generated by the ξ_i. Among others, this event guarantees that the number of roots in D is at most n^c, where c is a sufficiently small positive constant. [D was actually chosen so that the expectation of the number of roots in D is $O(1)$.] When $\mathcal{T}$ holds, we split $P=RQ$, where $R:=\prod_{\zeta_i\in D}(z-\zeta_i)$ and $Q:=\prod_{\zeta_i\notin D}(z-\zeta_i)$. Then

$$\begin{aligned}&\int_D\big|\log\big|P_{n,\xi}(z)\big|\Delta G(z)\big|^2\,dz\\&\quad\le 2\bigg(\int_D\big|\log\big|R_{n,\xi}(z)\big|\Delta G(z)\big|^2\,dz+\int_D\big|\log\big|Q_{n,\xi}(z)\big|\Delta G(z)\big|^2\,dz\bigg).\end{aligned}$$

The first integral on the RHS is easy to bound, as the number of roots in R is small, and $\log|R|$ can be split into sum of few terms. To bound the second one, we show that $|\log|Q_{n,\xi}(z)|\Delta G(z)|$ is small for *every point* in D. Typically, in order to prove that an event $\mathcal{E}(z)$ holds for every point z in some domain D one makes use of the ϵ-net argument. We put an ϵ-net on D and prove that $\mathcal{E}(z)$ holds for all points in the net, and then use some analytic argument to extend the net to the whole domain. If the net has size N, then by the union bound, we need to show that for each z in the net $\mathbf{P}(\mathcal{E}(z)\text{ holds}) \geq 1 - o(1/N)$. The proof of this usually requires sophisticated anti-concentration inequalities; furthermore, sometimes the bound itself is not true (which does not contradict the correctness of the final statement we want to prove). In our situation, we make a novel use of Harnack's inequality, which allows us to reduce the statement to one point, instead of to the whole ϵ-net, which completely avoids the use of union bound argument. This way, we obtain a sufficiently strong bound on the second moment so that the sampling procedure goes through. See Section 4.2 for more details.

The trickier part is when $\mathcal{T}$ does not hold. In this case, it is possible that sampling does not provide a good approximation. We are going to avoid this problem by directly showing that the contribution coming from the complement $\mathcal{T}^c$ of $\mathcal{T}$ toward the expectations in (3.1) is small, namely $\mathbf{E}_\xi F(Z)\mathbf{I}_{\mathcal{T}^c} = o(1)$ (and the same for the $\tilde{\xi}$ version).

The main difficulty here is that the logarithm function has a pole at zero. If $|P_{n,\xi}(z)|$ is very close to zero in some region, then the value of $\log|P_{n,\xi}(z)|$ could be very large. [Another type of danger is that $|P_{n,\xi}(z)|$ is large, but this is easy to deal with, even by elementary method such as the moment method.] To overcome this problem, one needs to show that with high probability, $|P_{n,\xi}(z)|$ is bounded away from 0. Technically speaking, we need to show

$$\mathbf{P}\big(|c_n\xi_n z^n + \cdots + c_0\xi_0| \leq \epsilon(n)\big)$$

is sufficiently small, for most value of z and a properly chosen parameter $\epsilon(n)$. These types of estimates are called anti-concentration (or small ball) inequality in the literature; see [31] for an introduction. This part is the most delicate part of our proof, and unlike prior works (see, e.g., [38]), our method could treat the general set of coefficients considered here.

In this paper, we introduce a completely different way to obtain the desired anti-concentration bound, which makes use of of various *a priori* estimates for $P_{n,\xi}$ and a recent powerful result of Nazarov–Nishry–Sodin [27] about the log-integrability of random Rademacher series. As a matter of fact, Nazarov et al. result only holds for random Rademacher variables (and may fail for others). We use a couple of symmetrization arguments to handle the general case. See Section 4 and in particular, Sections 4.1 and 4.4 for details.

By completing the above scheme, we obtain universality results for the complex roots. The handling of real roots also requires extra care. In order to prove the universality of the correlation functions among real roots (including the universality of the density function which yields new results on the expectation discussed in the Introduction) we need to show that there is no complex root near the real line, with high probability. This, at the intuition level at least, would allow us to translate results for complex roots near the real line to results for real roots, as once a root is sufficiently near the real line it has to be real.

One way to obtain this is via the so-called weak level repulsion property, relying on explicit estimates of the Kac–Rice formula for Kac polynomials with Gaussian coefficients. However, it is very difficult, if not impossible, to obtain similar estimates for the general polynomials considered in this paper, particularly in the case when the means of the coefficients are nonzero. We handle this problem by a novel argument, based on Rouché's theorem following an ideas from a paper of Peres and Virag [33] and the monograph by Hough et al. [13]. Apparently, the repulsion property is interesting on its own right, and there is a chance that the argument can be applied for other settings.

To illustrate the idea, let us consider a disk $B(x_0, r)$ center at a point x_0 on the real line. We want to show that if r is sufficiently small, then with high probability $B(x_0, r)$ contains at most one root. This excludes the complex roots as they come in conjugated pairs. Define $g(z) = P_{n,\xi}(x_0) + (z - x_0)P'_{n,\xi}(x_0)$. By Rouché's theorem, if we can show that (with high probability), $|P_{n,\xi}(z) - g(z)| < |g(z)|$ for all z on the boundary of $B(x_0, r)$, then $P_{n,\xi}(z)$ and $g(z)$ have the same number of roots inside the disk. Note that $g(z)$ is linear, so it has at most one root. The verification of $|P_{n,\xi}(z) - g(z)| < |g(z)|$ makes use of the Cauchy's integral formula and an anti-concentration result. (One can also use an ϵ-net argument here, but the details are more involved.) See Section 5 for details.

Finally, let us discuss the treatment of polynomials with Gaussian coefficients. The strategy of the proof of Theorem 1.4 (and other results in the Introduction) is to reduce to the Gaussian case, using universality results. In fact, the Gaussian setting of Theorem 1.4 and Theorem 1.8 are already substantially new, and furthermore our method of proof is novel compared to previous works. For example, the only case we know where the optimal error term $O(1)$ in our results was obtained is Kac polynomials, thanks to the very explicit formula (1.3). In our general setting while some version of (1.3) is available, evaluation of such formula turns out to be fairly delicate: in many other previous works for the mean-zero coefficients setting [6, 7, 35, 36] (see also [3]), researchers used the method of Logan and Shepp [23, 24], but this could not lead to the error term $O(1)$, and for coefficients with nonzero means (see below), the analysis from Farahmand's and Ibragimov–Maslova's paper [12, 17] do not lead to the error bound $O(1)$, even for the Kac polynomial. Finally, none of the above mentioned analysis can be reproduced to yield an asymptotic result for our general setting, where only the order of magnitude of the coefficients c_i is known.

Now let us discuss briefly the main new ideas in our the treatment of the Gaussian case. Via the Kac–Rice formula, the analysis of the nonzero mean case relies on several key estimates from the zero mean setting. In the zero mean case, our new idea is to develop a reformulation of the Edelman–Kostlan formula for the density function of the distribution of real zeros [10], so that the density could be computed using *only* the variance function $\mathbf{Var}[P_{n,\mathrm{Gauss}}]$ and its first few derivatives. This enables us to reduce the analysis of the density function to a careful study of the large n asymptotics of $\mathbf{Var}[P_{n,\mathrm{Gauss}}]$ and its derivatives. This novel approach allows us to get the $O(1)$ estimate for the error terms, which can not be obtained using the Logan–Shepp methods. The analysis of the large n behavior of $\mathbf{Var}[P_{n,\mathrm{Gauss}}]$ and its derivatives involves fairly technical estimates and occupies the last few sections of the paper. Unlike the Kac polynomials, in our setting the distribution of the real zeros is not invariant under the map $x \mapsto 1/x$, leading to extra difficulty in the analysis.

4. Proof of complex local universality for polynomials. Throughout the paper, $L := \frac{1}{\delta}$.

In this section, we prove Theorem 2.3. In particular, we will prove (2.8). The same proof works for (2.9) by replacing P by Q, unless otherwise noted. Notice that we only consider Q when talking about $\delta \geq \frac{1}{10n}$.

We can assume without loss of generality that $\tilde{\xi}_i$ has Gaussian distribution for all i.

By standard arguments using the Fourier analysis, using the assumption that the test function G is sufficiently smooth, one gets that G equals its Fourier series on its support with the Fourier coefficients growing sufficiently slowly. Therefore, if the desired statement is proven for each term (which is smoothly truncated on the support of G) in the Fourier expansion, it extends automatically to G. In other words, the problem reduces to proving (2.8) for

$$G(w_1, \dots, w_m) = G_1(w_1) \cdots G_k(w_k), \tag{4.1}$$

where for each $1 \leq i \leq k$, $G_i : \mathbb{C} \to \mathbb{C}$ is a smooth function supported in $B(0, 10^{-2})$ and $|\nabla^a G_i| \leq 1$ for all $0 \leq a \leq 3$.

When G is of that form, we have

$$\begin{aligned}
&\int_{\mathbb{C}^k} G(w_1, \dots, w_k) \rho^{(k)}_{\check{P}}(\check{z}_1 + w_1, \dots, \check{z}_k + w_k)\, dw_1 \cdots dw_k \\
&\quad = \mathbf{E} \sum_{i_1, \dots, i_k \text{ distinct}} G_1(\zeta^{\check{P}}_{i_1} - \check{z}_1) \cdots G_k(\zeta^{\check{P}}_{i_k} - \check{z}_k).
\end{aligned} \tag{4.2}$$

Let $r_0 = 10^{-2}$. By the inclusion-exclusion formula, we then can rewrite the later expression as

$$\mathbf{E} \prod_{j=1}^{k} X^P_j \tag{4.3}$$

plus a bounded number of lower order terms which are of the form (4.3) for smaller values of k, where

$$X^P_j = X^P_{\check{z}_j, G_j} = \sum_{i=1}^{n} G_j(\zeta^{\check{P}}_i - \check{z}_j). \tag{4.4}$$

Hence, by induction on k, it suffices to show that

$$\left|\mathbf{E}\prod_{j=1}^{k} X_j^P - \mathbf{E}\prod_{j=1}^{k} X_j^{\widetilde{P}}\right| \le C\delta^c. \tag{4.5}$$

If P does not vanish on the support of H_j, then by the Green formula it follows from (4.4) that

$$X_j^P = \int_{\mathbb{C}} \log\bigl|\check{P}(z)\bigr| H_j(z)\,dz = \int_{B(\check{z}_j, r_0)} \log\bigl|\check{P}(z)\bigr| H_j(z)\,dz, \tag{4.6}$$

where $H_j(z) = -\frac{1}{2\pi}\triangle G_j(z - \check{z}_j)$. Note that $\operatorname{supp}(H_j) \subset B(\check{z}_j, r_0)$.

Let $K_j^P = \log|\check{P}(z)|H_j(z)$. Let c_1 be a small positive constant to be chosen later. Let $\mathcal{T} = \mathcal{T}(\delta)$ be the event on which:

(i) $P \not\equiv 0$.
(ii) $N_P(B(z_j, \frac{\delta}{10})) \le L^{c_1}$ for all $1 \le j \le k$.
(iii) $\log|P(z_j)| \ge -\frac{1}{2}L^{c_1}$ for all $1 \le j \le k$.
(iv) $\log|P(z)| \le \frac{1}{2}L^{c_1}$ for all z such that $|z| \in I(\delta) + (-\delta/2, \delta/2)$.

And if $\delta \ge \frac{1}{10n}$, we also require that on the event $\mathcal{T}$:

(v) $N_Q(B(z_j, \frac{\delta}{10})) \le L^{c_1}$ for all $1 \le j \le k$.
(vi) $\log|Q(z_j)| \ge -\frac{1}{2}L^{c_1}$ for all $1 \le j \le k$.
(vii) $\log|Q(z)| \le \frac{1}{2}L^{c_1}$ for all z such that $|z| \in I(\delta) + (-\delta/2, \delta/2)$.

The rest of the proof consists of several parts. In Section 4.1, we will show that the event $\mathcal{T}$ occurs with high probability. Then in Section 4.2, we will show that $\|K_j^P\|_{L^2(z)}$ is small on $\mathcal{T}$ for all $1 \le j \le k$. This allows us to approximate X_j^P by a finite sum $\frac{1}{m}\sum_{i=1}^m \log|\check{P}(w_i)|H_j(w_i)$ using the Monte Carlo sampling method. After the approximation step, in Section 4.3, we show that the two approximating expressions for P and $\widetilde{P}$ are close using the Lindeberg swapping technique. Next, in Section 4.4, we show that the tail event $\mathcal{T}^c$ does not contribute significantly to the picture, that is, $\mathbf{E}(|\prod_{j=1}^k X_j^P|\mathbf{1}_{\mathcal{T}^c})$ is small. This is the key step of our proof. Finally, we wrap up the proof in Section 4.5.

4.1. *The event $\mathcal{T}$ occurs with high probability.* Let A be a large constant, say $A = k + 2$. And set

$$\gamma(\delta) = \begin{cases} \delta^A & \text{if } \dfrac{\log^2 n}{n} \le \delta \le \dfrac{1}{C}, \\ n^{-1/2} & \text{if } 0 \le \delta < \dfrac{\log^2 n}{n}. \end{cases}$$

In this section, we show that $\mathbf{P}(\mathcal{T}) \ge 1 - C\gamma(\delta)$ for some constant C. To show that (iii) and (vi) occur with high probability, we will need two Littlewood–Offord-type anti-concentration bounds. The first bound for ξ_i being Rademacher is known as Erdős' lemma. We reduce the general case to the Rademacher case and then include a proof of the Erdős' lemma.

LEMMA 4.1. *If the ξ_i's satisfy Condition 1, there exists a constant D such that for any integer $n \ge 1$, real number $a > 0$, and complex numbers $a_1, \dots, a_n$ with $|a_i| \ge a$ for all i, and for any $z \in \mathbb{C}$, we have*

$$\mathbf{P}\left(\left|\sum_{i=1}^n a_i\xi_i - z\right| \le \frac{a}{D}\right) \le \frac{D}{\sqrt{n}}.$$

PROOF OF LEMMA 4.1. By translation, we can assume that $\mathbf{E}\xi_i = 0$ for all i. It then suffices to show the lemma when the ξ_i's and a_i's are real. Indeed, assume that the statement on the real line holds true. In the general case, assume without loss of generality (wlog) that $a = 1$. Since $|a_i| \ge 1$, either $|\mathrm{Re}(a_i)| \ge \max\{|\mathrm{Im}(a_i)|, \frac{1}{\sqrt{2}}\}$ or $|\mathrm{Im}(a_i)| \ge \max\{|\mathrm{Re}(a_i)|, \frac{1}{\sqrt{2}}\}$. By the pigeonhole principle, we can assume wlog that there are at least $n/2$ indices i such that $|\mathrm{Re}(a_i)| \ge \max\{|\mathrm{Im}(a_i)|, \frac{1}{\sqrt{2}}\}$. For such i, set $X_i = \mathrm{Re}(\xi_i) - \frac{\mathrm{Im}(a_i)}{\mathrm{Re}(a_i)}\mathrm{Im}(\xi_i)$, $Y_i = \mathrm{Im}(\xi_i) + \frac{\mathrm{Im}(a_i)}{\mathrm{Re}(a_i)}\mathrm{Re}(\xi_i)$, then $a_i\xi_i = \mathrm{Re}(a_i)(X_i + \sqrt{-1}Y_i)$, $\mathbf{Var}(X_i) + \mathbf{Var}(Y_i) = 1 + \frac{\mathrm{Im}^2(a_i)}{\mathrm{Re}^2(a_i)} \in [1, 2]$, and $\mathbf{E}|X_i|^{2+\epsilon}, \mathbf{E}|Y_i|^{2+\epsilon} \le 2^{2+\epsilon}\mathbf{E}|\xi|^{2+\epsilon} \le 2^{2+\epsilon}\tau_2$. By the pigeonhole principle, we can then assume wlog that there are at least $n/4$ indices i such that $|\mathrm{Re}(a_i)| \ge 1/\sqrt{2}$ and $\mathbf{Var}(Y_i) \in [1/2, 2]$. Now, for such i, $\mathrm{Re}(a_i)Y_i = \mathrm{Re}(a_i)\sqrt{\mathbf{Var}(Y_i)}\frac{Y_i}{\sqrt{\mathbf{Var}(Y_i)}}$ with $|\mathrm{Re}(a_i)|\sqrt{\mathbf{Var}(Y_i)} \ge \frac{1}{2}$. This allows one to use the result for the reals (after conditioning on the rest Y_j's) with coefficients $\mathrm{Re}(a_i)\sqrt{\mathbf{Var}(Y_i)}$ and random variables $\frac{Y_i}{\sqrt{\mathbf{Var}(Y_i)}}$ and obtain a constant D such that for any $y \in \mathbb{R}$,

$$\mathbf{P}\left(\left|\sum_{i=1}^n \mathrm{Re}(a_i)Y_i - y\right| \le \frac{1}{2D}\right) \le \frac{D}{\sqrt{n}}.$$

This implies that for all $z \in \mathbb{C}$, $\mathbf{P}(|\sum_{i=1}^n a_i\xi_i - z| \le \frac{1}{2D}) \le \frac{D}{\sqrt{n}} \le \frac{2D}{\sqrt{n}}$.

Thus, we can assume that the ξ_i's and a_i's are real. We can further assume that the a_i's have the same sign. Indeed, by the pigeonhole principle again, there are at least $n/2$ numbers a_i having the same sign, say, positive. By conditioning on the ξ_i's with a_i negative, we can reduce the problem to the case $a_i > 0$ for all i. Thus, the assumption becomes $a_i \ge 1$, $\forall i$.

Since the ξ_i's satisfy Condition 1, there exist constants D and $q > 0$ such that $\mathbf{P}(D \ge \xi_i - \xi_i' \ge \frac{1}{D}) \ge q$, where ξ_i' is an independent copy of ξ_i. Let $\epsilon_1, \dots, \epsilon_n$ be independent Rademacher random variables which are independent of all previous random variables. Let

$$\tilde{\xi}_i = \begin{cases} \xi_i & \text{if } \epsilon_i = 1, \\ \xi_i' & \text{if } \epsilon_i = -1. \end{cases}$$

Then $\tilde{\xi}_1, \dots, \tilde{\xi}_n$ are independent random variables having the same distribution as $\xi_1, \dots, \xi_n$. Hence, it suffices to show that for all $x \in \mathbb{R}$,

$$\mathbf{P}\left(\left|\sum_{i=1}^n a_i \tilde{\xi}_i - x\right| \le \frac{1}{3D}\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

Let J be the set of indices j such that $\xi_j - \xi_j' \ge \frac{1}{D}$. Since $\mathbf{P}(\xi_j - \xi_j' \ge \frac{1}{D}) \ge q$, $\mathbf{E}|J| \ge nq$. By Chernoff's bound (see, for instance, [9], Theorem 1.1),

$$\mathbf{P}\left(|J| \le \frac{nq}{2}\right) \le \mathbf{P}\left(|J| \le \frac{\mathbf{E}|J|}{2}\right) \le 2e^{-\frac{\mathbf{E}|J|}{8}} \le 2e^{-\frac{nq}{8}} \le \frac{1}{\sqrt{n}}. \tag{4.7}$$

Conditioning on the event that $|J| \ge \frac{nq}{2}$, and fixing $\tilde{\xi}_k$'s for all $k \notin J$ as well as ξ_j's, ξ_j''s for all $j \in J$, the only source of randomness left is from ϵ_j's with $j \in J$. It suffices to show that for all x, $\mathbf{P}(|\sum_{j \in J} a_j \tilde{\xi}_j - x| \le \frac{1}{3D}) = O(\frac{1}{\sqrt{n}})$.

Let $\mathcal{F}$ be the collection of all subsets $\{j \in J : \epsilon_j = 1\}$ as ϵ_j run over all possible values such that $|\sum_{j \in J} a_j \tilde{\xi}_j - x| \le \frac{1}{3D}$. Observe that $\mathcal{F}$ is an anti-chain. Indeed, suppose that $F \subset F'$ be two elements of $\mathcal{F}$ which correspond to $\epsilon_j = x_j$ and $\epsilon_j = x_j'$, respectively, $(x_j, x_j' \in \{\pm 1\})$. For $\epsilon_j = x_j$,

$$\sum_{j \in J} a_j \tilde{\xi}_j = \sum_{j \in F} a_j \xi_j + \sum_{j \in J \setminus F} a_j {\xi'}_j, \tag{4.8}$$

and for $\epsilon_j = x_j'$,

$$\sum_{j \in J} a_j \tilde{\xi}_j = \sum_{j \in F'} a_j \xi_j + \sum_{j \in J \setminus F'} a_j {\xi'}_j, \tag{4.9}$$

The difference of the expressions in (4.8) and (4.9) is

$$\sum_{j \in F' \setminus F} a_j (\xi_j - \xi_j') \ge \frac{1}{D}$$

which contradicts the assumption that they both lie in an interval of length at most $\frac{2}{3D}$. Hence, $\mathcal{F}$ is an anti-chain. And so, $|\mathcal{F}| \le \binom{|J|}{\lfloor |J|/2 \rfloor}$ by Sperner's theorem [1], Chapter 12. It follows that of all $2^{|J|}$ choices of the values of ϵ_j, there are at most $\binom{|J|}{\lfloor |J|/2 \rfloor}$ of them can make $|\sum_{j \in J} a_j \tilde{\xi}_j - x| \le \frac{1}{3D}$. By Stirling's formula,

$$\mathbf{P}\left(\left|\sum_{j \in J} a_j \tilde{\xi}_j - x\right| \le \frac{1}{3D}\right) \le \frac{\binom{|J|}{\lfloor |J|/2 \rfloor}}{2^{|J|}} = O\left(\frac{1}{\sqrt{|J|}}\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

This completes the proof. $\square$

The next bound is proven in [38], Lemma 9.2. We include a short proof for the convenience of the reader.

LEMMA 4.2. *Let $(\xi_i)_{i=1}^n$ be independent random variables satisfying Condition 1. There exist positive constants C' and α such that for any complex number z, any integer m, and any sequence of complex numbers $e_0, \dots, e_n$ containing a lacunary subsequence $|e_{i_1}| \geq 2|e_{i_2}| \geq \dots \geq 2^m|e_{i_m}|$, we have*

$$\mathbf{P}\left(\left|\sum_{i=0}^n e_i\xi_i - z\right| \leq |e_{i_m}|\right) \leq C'\exp(-\alpha m). \tag{4.10}$$

PROOF OF LEMMA 4.2. As in the proof of Lemma 4.1, we can assume that $\mathbf{E}\xi = 0$. Consider ξ_i', $\tilde{\xi}_i$, D and q as in that proof. Without loss of generality, assume that $D \geq 10$. We can choose a subsubsequence $(e_{i_{j_k}})_{k=1}^{\tilde{m}}$ of the lacunary sequence $(e_{i_j})_{j=1}^m$ with $\tilde{m} = \Theta(m)$ such that $|e_{i_{j_1}}| \geq D^3|e_{i_{j_2}}| \geq \dots \geq D^{3\tilde{m}}|e_{i_{j_k}}|$. By conditioning on the random variables ξ_l with l not equal any i_{j_k}, we can assume that the subsubsequence equals the original sequence; in other words, $i_{j_k} = k$ for all k, and $\tilde{m} = m = n$. Let J be the set of indices $j < m$ such that $D \geq \xi_j - \xi_j' \geq \frac{1}{D}$. By the same argument with Chernoff's bound as before, we have $|J| \geq \frac{mq}{2}$ with probability at least $1 - \exp(-\alpha m)$. Conditioning on the event that $|J| \geq \frac{mq}{2}$, and fixing $\tilde{\xi}_k$'s for all $k \notin J$ as well as ξ_j's, ξ_j''s for all $j \in J$, the only source of randomness left is from ϵ_j's with $j \in J$. It suffices to show that for all z, $\mathbf{P}(|\sum_{j\in J} e_j\tilde{\xi}_j - z| \leq |e_m|) = O(\exp(-\alpha m))$. By triangle inequality, we can show that for any two instances of $(\epsilon_j)_{j\in J}$, the difference of the two sums $\sum_{j\in J} e_j\tilde{\xi}$ has magnitude at least $4|e_m|$. And so, $\mathbf{P}(|\sum_{j\in J} e_j\tilde{\xi}_j - z| \leq |e_m|) \leq 2^{-|J|} = O(\exp(-\alpha m))$. $\square$

We are now ready to show that (iii) and (vi) occur with the desired probability. For $\delta \in [\frac{\log^2 n}{n}, \frac{1}{C}]$, we prove the following.

LEMMA 4.3. *For any constants $A > 0$ and $c > 0$, there exists a constant C such that for any $\delta \in [\frac{\log^2 n}{n}, \frac{1}{C}]$, complex number z such that $|z| \in I(\delta)$, and $1 \leq \lambda \leq \frac{n\delta}{\log^2 \delta}$, one has*

$$\mathbf{P}\left(\log|P(z)| \geq -\frac{1}{2}\lambda\delta^{-c}\right) \geq 1 - C\frac{\delta^A}{\lambda^A}, \tag{4.11}$$

$$\mathbf{P}\left(\log|Q(z)| \geq -\frac{1}{2}\lambda\delta^{-c}\right) \geq 1 - C\frac{\delta^A}{\lambda^A}. \tag{4.12}$$

PROOF. Since $L = \frac{1}{\delta} \leq \frac{n}{\log^2 n}$, we have $L\log^2 L \leq \frac{n}{\log^2 n}\log^2 n = n$. Thus, there exists some λ such that $1 \leq \lambda \leq \frac{n}{L\log^2 L} = \frac{n\delta}{\log^2 \delta}$. Set $m = \lceil \frac{\log C' + A\log(\lambda L)}{\alpha} \rceil$, then $C'\exp(-\alpha m) \leq \frac{1}{\lambda^A L^A}$. We obtain a lacunary sequence $|c_{j_0}z^{j_0}| \geq 2|c_{2j_0}z^{2j_0}| \geq \dots \geq 2^m|c_{i_0}z^{i_0}|$ where $j_0 = \lceil BL \rceil$, B is a large enough constant and $i_0 = (m+1)j_0$.

Observe that $i_0 \le \frac{n}{2}$ and $c_{i_0}|z|^{i_0} \ge e^{-1/2\lambda L^c}$. Thus, by applying inequality (4.10) to this lacunary sequence, we get (4.11).

To prove (4.12), we similarly apply (4.10) to the lacunary sequence

$$\left|\frac{d_{j_0}}{d_0} z^{j_0}\right| \ge 2\left|\frac{d_{2j_0}}{d_0} z^{2j_0}\right| \ge \cdots \ge 2^m \left|\frac{d_{i_0}}{d_0} z^{i_0}\right|. \qquad \square$$

For $\delta \in [\frac{1}{20n}, \frac{\log^2 n}{n}]$, we prove the following.

LEMMA 4.4. *For any positive constant c, there exists a constant C such that for all $\delta \in [\frac{1}{20n}, \frac{\log^2 n}{n}]$, and complex number z such that $|z| \in I(\delta)$, it holds that $\log|P(z)| \ge -\frac{1}{2}\delta^{-c}$ with probability at least $1 - Cn^{-1/2}$.*

If $\delta \ge \frac{1}{10n}$, the same statement holds for Q in place of P.

PROOF. If $\delta \in [\frac{1}{10n}, \frac{\log^2 n}{n}]$, $|z| \in [1-2\delta, 1-\delta]$, and $N_0 \le i \le n$, then

$$|c_i z^i| \ge \tau_1 n^{-|\rho|}(1-2\delta)^n \ge \tau_1 n^{-|\rho|}\left(1 - \frac{2\log^2 n}{n}\right)^n \ge e^{-8\log^2 n} \ge 2De^{-\frac{1}{2}L^c},$$

where D is the constant in Lemma 4.1.

By Lemma 4.1, we have $\mathbf{P}(|P(z)| \le e^{-\frac{1}{2}L^c}) \le Cn^{-1/2}$. Note that we may not have $|c_i z^i| \ge 2De^{-\frac{1}{2}L^c}$ for $i < N_0$, but by first conditioning on $\xi_0, \dots, \xi_{N_0}$, Lemma 4.1 still gives us the desired result.

The same argument holds for $\delta \le \frac{1}{10n}$ and for Q in place of P. $\square$

In the following lemma, we show that the events (iv) and (vii) occur with high probability.

LEMMA 4.5. *For any constants $A > 1$ and $c > 0$, there exists a constant C such that for any $\frac{1}{10n} \le \delta \le \frac{1}{C}$ and $\lambda \ge 1$, we have*

$$\log M \le \frac{1}{2}\lambda L^c$$

with probability at least $1 - \frac{\delta^A}{\lambda^A}$, where $M = \max\{|P(z)|, |Q(z)| : |z| \le 1 - \delta/2\}$.

And if $\frac{1}{20n} \le \delta \le \frac{1}{10n}$ then

$$\log M \le \frac{1}{2}L^c$$

with probability at least $1 - n^{-1/2}$, where $M = \max\{|P(z)| : |z| \le 1 + \frac{4}{n}\}$.

PROOF. Assume that $\frac{1}{10n} \le \delta \le \frac{1}{C}$. Let $X = \frac{2-\delta/2}{2-\delta} \in (1, 2)$ and $a_i = \lambda^A L^A X^i$. Let

$$\Omega' = \{\omega : |\xi_i| \le a_i, \forall i = 0, \dots, n\}.$$

The probability of the complement of Ω' is

$$\mathbf{P}(\Omega'^c) = \mathbf{P}(\exists i \in \{0, 1, \dots, n\} : |\xi_i| > a_i) \le \sum_{i=0}^{n} \frac{\tau_2}{a_i^2} \le \frac{1}{\lambda^A L^A}.$$

For every $\omega \in \Omega'$, we have

$$\begin{aligned}\max_{|z|\le 1-\delta/2} |P(z)| &\le \sum_{i=0}^{n} |c_i \xi_i| \left(1 - \frac{\delta}{2}\right)^i \le \sum_{i=0}^{n} a_i |c_i| \left(1 - \frac{\delta}{2}\right)^i \\ &\le C'' \lambda^A L^A \left(\frac{4}{\delta}\right)^{\lceil \rho \rceil + 1} \le e^{\frac{1}{2}\lambda L^c}.\end{aligned}$$

A similar bound holds for Q.

When $\frac{1}{20n} \le \delta \le \frac{1}{10n}$, we set $\Omega' = \{\omega : |\xi_i| \le n, \forall 0 \le i \le n\}$ and argue similarly. □

Combining Lemmas 4.3, 4.4 and 4.5, we obtain that the events (ii) and (v) occur with high probability.

PROPOSITION 4.6 (Nonclustering). *For any constants $A > 1$ and $c > 0$, there exists a constant C such that*:

(i) *For any $\delta \in [\frac{\log^2 n}{n}, \frac{1}{C}]$, $1 \le \lambda \le \frac{n\delta}{\log^2 \delta}$, and complex number z such that $|z| \in I(\delta)$, we have*

$$N_P\big(B(z, \delta/9)\big) \le \lambda \delta^{-c} \quad \textit{and} \quad N_Q\big(B(z, \delta/9)\big) \le \lambda \delta^{-c}$$

with probability at least $1 - C\frac{\delta^A}{\lambda^A}$.

(ii) *For any $\delta \in [\frac{1}{20n}, \frac{\log^2 n}{n}]$ and complex number z such that $|z| \in I(\delta)$, one has*

$$N_P\big(B(z, \delta/9)\big) \le \delta^{-c}$$

with probability at least $1 - Cn^{-1/2}$. The same statement holds for Q in place of P when $\delta \in [\frac{1}{10n}, \frac{\log^2 n}{n}]$.

PROOF. By our convention, we only need to work on the event that P and Q do not vanish identically. In the following, we prove for P. The same argument works for Q equally well.

We first prove (i). By Jensen's inequality, we have

$$N_P\big(B(z, s)\big) \le \frac{\log M - \log |P(z)|}{\log \frac{R}{s}} \le \log M - \log\big|P(z)\big|,$$

where $R = \frac{\delta}{3}$, $s = \frac{\delta}{9}$, $M = \max_{|w-z|=R} |P(w)| \le \max_{|w|\le 1-\delta/2} |P(w)|$.

Claim (i) follows from Lemmas 4.3 and 4.5. Similarly, (ii) follows from Lemmas 4.4 and 4.5. □

From the above proposition, we obtain the following.

PROPOSITION 4.7. *For any constants $A > 1$ and $c_1 > 0$, there exists a constant C such that for any $\frac{1}{20n} \le \delta \le \frac{1}{C}$, we have*

$$\mathbf{P}(\mathcal{T}(\delta)) \ge 1 - C\gamma(\delta).$$

PROOF. By Hölder's inequality,

$$\begin{aligned} 1 = \mathbf{Var}(\xi_i) &\le \mathbf{E}|\xi_i|^2 \le (\mathbf{E}|\xi_i|^{2+\epsilon})^{2/(2+\epsilon)} \mathbf{P}(|\xi_i| > 0)^{\epsilon/(2+\epsilon)} \\ &\le \tau_4^{2/(2+\epsilon)} \mathbf{P}(|\xi_i| > 0)^{\epsilon/(2+\epsilon)}. \end{aligned}$$

Thus, for all i, $\mathbf{P}(\xi_i = 0) \le 1 - \frac{1}{C'}$ for some constant C'. This gives

$$\mathbf{P}(P \equiv 0) = \left(1 - \frac{1}{C}\right)^n \le Cn^{-A} \le C\gamma(\delta).$$

The proposition then follows from Lemmas 4.3, 4.4, 4.5, 4.6 and the union bound. □

4.2. *Approximation of integrals by finite sums*. Fix $\delta \in [\frac{1}{20n}, \frac{1}{C}]$. In this section, we show that on the event $\mathcal{T}$, the norms $\|K_j^P\|_{L^2(z)}$ are small for all $1 \le j \le k$. At the end of the section, this bound allows us to use the Monte Carlo sampling lemma to approximate X_j^P with finite (sample) sums, on which we will apply the Lindeberg swapping argument. The crucial tool in this section is Harnack's inequality which allows us to show that property (iii) in the definition of $\mathcal{T}$ basically holds for every $z \in B(z_j, 10^{-5}\delta)$.

Recall that $K_j^P = \log|\check{P}(z)|H_j(z)$. and

$$\begin{aligned} \|K_j^P\|_{L^2(z)}^2 &= \int_{B(\check{z}_j, r_0)} \left|\log|\check{P}(z)|H_j(z)\right|^2 dz \\ &\le \int_{B(\check{z}_j, r_0)} \left|\log|\check{P}(z)|\right|^2 dz \\ &= \frac{1}{10^{-6}\delta^2} \int_{B(z_j, 10^{-5}\delta)} \log^2|P(z)|\, dz. \end{aligned} \tag{4.13}$$

LEMMA 4.8. *On $\mathcal{T}$, one has the bound*:

$$\left\|\log|P(z)|\right\|_{L^2(B(z_j, 10^{-5}\delta))} \le L^{4c_1 - 1}. \tag{4.14}$$

Note that this is a deterministic statement.

PROOF. Fix $\omega \in \mathcal{T}$. Consider $I := [10^{-5}\delta, 10^{-1}\delta]$, we have $|I| \geq \frac{\delta}{20}$. There exists an $r \in I$ such that P does not have zeros in the (closed) annulus $A(z_j, r - \eta, r + \eta)$ where $\eta = \frac{1}{80}\delta^{1+c_1}$. Indeed, assume such an r does not exist, then

$$N_P B(z_j, \delta/10) \geq \frac{|I|}{3\eta} > \delta^{-c_1}$$

which contradicts the condition (ii) in the definition of $\mathcal{T}$.

Now fix that r, we have $\int_{B(z_j,10^{-5}\delta)} \log^2 |P(z)|\,dz \leq \int_{B(z_j,r)} \log^2 |P(z)|\,dz$. Let $\zeta_1, \dots, \zeta_m$ be all zeros of P in $B(z_j, r - \eta)$, then $m \leq L^{c_1}$ and $P(z) = (z - \zeta_1)\cdots(z - \zeta_m)g(z)$ where g is a polynomial having no zeros on the closed ball $B(z_j, r + \eta)$. We have

$$\begin{aligned}\big\|\log\big|P(z)\big|\big\|_{L^2(B(z_j,r))} &\leq \sum_{i=1}^{m} \big\|\log|z - \zeta_i|\big\|_{L^2(B(z_j,r))} + \big\|\log\big|g(z)\big|\big\|_{L^2(B(z_j,r))} \\ &\leq m\delta^{1-c_1} + \big\|\log\big|g(z)\big|\big\|_{L^2(B(z_j,r))},\end{aligned}$$

where the last inequality is because

$$\int_{B(z_j,r)} \log^2 |z - \zeta_i|\,dz \leq \int_{B(0,\delta)} \log^2 |z|\,dz \leq \delta^{2-2c_1}.$$

Thus,

$$\big\|\log\big|P(z)\big|\big\|_{L^2(B(z_j,r))} \leq L^{2c_1-1} + \big\|\log\big|g(z)\big|\big\|_{L^2(B(z_j,r))}. \tag{4.15}$$

Next, we will estimate $\int_{B(z_j,r)} \log^2 |g(z)|\,dz$. Since $\log|g(z)|$ is harmonic in $B(z_j, r)$, it attains its extrema on the boundary. Thus,

$$\big\|\log\big|g(z)\big|\big\|_{L^2(B(z_j,r))} = \left(\int_{B(z_j,r)} \log^2\big|g(z)\big|\,dz\right)^{1/2} \leq \delta \max_{z\in\partial B(z_j,r)} \big|\log\big|g(z)\big|\big|. \tag{4.16}$$

Notice that $\log|g(z)|$ is also harmonic on the ball $B(z_j, r + \eta)$.

CLAIM 4.9. *For every z in $B(z_j, r + \eta)$, we have*

$$\log\big|g(z)\big| \leq L^{2c_1}.$$

PROOF. Since a harmonic function attains its extrema on the boundary, we can assume that $z \in \partial B(z_j, r + \eta)$. Since $|z| < |z_j| + \delta/2$, $|z| \in I(\delta) + (-\delta/2, \delta/2)$. So, by condition (iv) in the definition of $\mathcal{T}$, $\log|P(z)| \leq L^{c_1}$. Additionally, by noticing that $|z - \zeta_i| \geq 2\eta$ for all $1 \leq i \leq m$, we get

$$\log\big|g(z)\big| = \log\big|P(z)\big| - \sum_{i=1}^{m} \log|z - \zeta_i| \leq L^{c_1} - m\log(2\eta) \leq L^{2c_1}$$

as desired. $\square$

Now let $u(z) = L^{2c_1} - \log|g(z)|$, then u is a nonnegative harmonic function on the ball $B(z_j, r + \eta)$. By Harnack's inequality (see [34], Chapter 11) for the subset $B(z_j, r)$ of the above ball, we have that for every $z \in B(z_j, r)$,

$$\alpha u(z_j) \le u(z) \le \frac{1}{\alpha} u(z_j),$$

where $\alpha = \frac{\eta}{2r+\eta} \ge \frac{\delta^{c_1}}{160}$. Hence,

$$\alpha\bigl(L^{2c_1} - \log\bigl|g(z_j)\bigr|\bigr) \le L^{2c_1} - \log\bigl|g(z)\bigr| \le \frac{1}{\alpha}\bigl(L^{2c_1} - \log\bigl|g(z_j)\bigr|\bigr).$$

And so,

$$\bigl|\log\bigl|g(z)\bigr|\bigr| \le \frac{1}{\alpha}\bigl|\log\bigl|g(z_j)\bigr|\bigr| + \frac{1}{\alpha} L^{2c_1} \le 160 L^{c_1}\bigl|\log\bigl|g(z_j)\bigr|\bigr| + 160 L^{3c_1}. \tag{4.17}$$

Thus, we reduce the problem to bounding $|\log|g(z_j)||$. From Claim 4.9 and the condition (iii) in the definition of $\mathcal{T}$, we have

$$L^{2c_1} \ge \log\bigl|g(z_j)\bigr| = \log\bigl|P(z_j)\bigr| - \sum_{i=1}^{m} \log|z_j - \zeta_i| \ge \log\bigl|P(z_j)\bigr| \ge -\frac{1}{2} L^{c_1}.$$

And so, $|\log|g(z_j)|| \le L^{2c_1}$, which together with (4.17) give

$$\bigl|\log\bigl|g(z)\bigr|\bigr| \le 320 L^{3c_1}. \tag{4.18}$$

From (4.15), (4.16) and (4.18), we obtain

$$\bigl\|\log\bigl|P(z)\bigr|\bigr\|_{L^2(B(z_j,r))} \le L^{4c_1 - 1}.$$

Lemma 4.8 is proved. □

From this lemma, we conclude that on the event $\mathcal{T}$,

$$\|K_j^P\|_{L^2(z)} \le 10^3 \cdot L\bigl\|\log\bigl|P(z)\bigr|\bigr\|_{L^2(B(z_j, 10^{-5}\delta))} \le 10^6 L^{4c_1}. \tag{4.19}$$

Having bounded the 2-norm, we now use the following sampling lemma.

LEMMA 4.10 (Monte Carlo sampling lemma ([39], Lemma 38)). *Let (X, μ) be a probability space, and $F : X \to \mathbb{C}$ be a square integrable function. Let $m \ge 1$, let $x_1, \dots, x_m$ be drawn independently at random from X with distribution μ, and let S be the empirical average*

$$S := \frac{1}{m}\bigl(F(x_1) + \cdots + F(x_m)\bigr).$$

Then S has mean $\int_X F\,d\mu$ and variance $\frac{1}{m}\int_X (F - \int_X F\,d\mu)^2\,d\mu$. In particular, by Chebyshev's inequality, we have

$$\mathbf{P}\left(\left|S - \int_X F\,d\mu\right| \ge \lambda\right) \le \frac{1}{m\lambda^2}\int_X \left(F - \int_X F\,d\mu\right)^2 d\mu.$$

Conditioning on $\mathcal{T}$ and applying this sampling lemma, we have for large $m_0 > 0$ and small $\gamma_0 > 0$ to be chosen later,

$$\left| X_j^P - \frac{\pi r_0^2}{m_0} \sum_{i=1}^{m_0} K_j^P(\check{w}_{j,i}) \right| \le \frac{2\sqrt{\pi r_0^2}}{\sqrt{m_0 \gamma_0}} 10^6 L^{4c_1} \le \frac{C L^{4c_1}}{\sqrt{m_0 \gamma_0}} \tag{4.20}$$

with probability at least $1 - \gamma_0$, where $\check{w}_{j,i}$ are chosen independently at random from $B(\check{z}_j, r_0)$ with uniform distribution and are independent from all previous random variables. By exactly the same argument, (4.19) also holds for Q when $\frac{1}{10n} \le \delta \le \frac{1}{C}$.

4.3. *Log comparability.* We shall show in Section 4.5 that (4.20) allows us to reduce the problem to comparing $F(\log|P(z_1)|, \dots, \log|P(z_m)|)$ and $F(\log|\widetilde{P}(z_1)|, \dots, \log|\widetilde{P}(z_m)|)$ for some smooth function F. This is done by making use of the beautiful Lindeberg swapping trick. The following result is from [38]; we include a proof in the Appendix for the reader's convenience.

THEOREM 4.11 (Comparability of log-magnitude). *Let P be the random polynomial of the form (2.4) satisfying Condition 1(1). And let $\widetilde{P} = \sum_{i=0}^n c_i \tilde{\xi}_i z^i$ be the corresponding polynomial with Gaussian random variables $\tilde{\xi}_i$. Assume that $\tilde{\xi}_i$ matches moments to second order with ξ_i for every $i \in \{0, \dots, n\} \setminus I_0$ for some (deterministic) set I_0 (may depend on n) of size at most N_0 and that $\sup_{i\ge 0} \mathbf{E}|\tilde{\xi}_i|^{2+\epsilon} \le \tau_2$ where N_0 and τ_2 are constants in Condition 1(1).*

Then there exists a constant C_2 such that the following holds true. Let $\alpha_1 \ge C_2\alpha_0 > 0$ and $C > 0$ be any constants. Let $\delta \in (0, 1)$ and $m \le \delta^{-\alpha_0}$ and $z_1, \dots, z_m \in \mathbb{C}$ be complex numbers such that

$$\frac{|c_i||z_j|^i}{\sqrt{V(z_j)}} \le C\delta^{\alpha_1} \qquad \forall i = 0, \dots, n, j = 1, \dots, m, \tag{4.21}$$

where $V(z_j) = \sum_{i \in \{0,\dots,n\}\setminus I_0} |c_i|^2 |z_j|^{2i}$.

Let $F : \mathbb{C}^m \to \mathbb{C}$ be any smooth function such that $|\nabla^a F(w)| \le C\delta^{-\alpha_0}$ for all $0 \le a \le 3$ and $w \in \mathbb{C}^m$, then

$$\left|\mathbf{E}F\left(\log\left|P(z_1)\right|, \dots, \log\left|P(z_m)\right|\right) - \mathbf{E}F\left(\log\left|\widetilde{P}(z_1)\right|, \dots, \log\left|\widetilde{P}(z_m)\right|\right)\right| \le \widetilde{C}\delta^{\alpha_0},$$

where $\widetilde{C}$ is a constant depending only on α_0, α_1, C and not on δ.

Now we show that condition (4.21) holds for P and Q.

LEMMA 4.12. *Under the assumptions Theorem 2.3, there exist constants $\alpha_1 > 0$ and $C > 0$ such that for every $\delta \in [\frac{1}{20n}, \frac{1}{C}]$ and for every z such that $|z| \in I(\delta) + [-\delta/2, \delta/2]$,*

$$\frac{|c_i||z|^i}{\sqrt{\mathbf{Var}\, P(z)}} \le C\delta^{\alpha_1} \qquad \forall 0 \le i \le n. \tag{4.22}$$

and if $\delta \in [\frac{1}{10n}, \frac{1}{C}]$,

$$\frac{\frac{|d_i|}{|d_0|}|z|^i}{\sqrt{\mathbf{Var}\, Q(z)}} \le C\delta^{\alpha_1} \qquad \forall 0 \le i \le n. \tag{4.23}$$

Notice that once (4.22) holds, say, the contribution of a few terms in $\mathbf{Var}\, P(z) = \sum_{i=0}^{n} |c_i|^2 |z|^{2i}$ is negligible, and hence for any set I_0 of size at most N_0, we have $\frac{|c_i||z|^i}{\sqrt{\sum_{i\in\{0,\dots,n\}\setminus I_0} |c_i|^2|z|^{2i}}} \le C\delta^{\alpha_1}$ as required in (4.21).

PROOF. Let $\alpha_1 = \min(\rho + 1/2, 1/2) > 0$. We prove (4.22) when $\delta \in [\frac{1}{20n}, \frac{1}{C}]$. The other parts of the statement are similar. Recall that $L \le 20n$. We have from (2.5)

$$\mathbf{Var}\, P(z) = \sum_{i=0}^{n} c_i^2 |z|^{2i} \ge \frac{\tau_1^2}{40^{2\rho}} \sum_{i=\lfloor L/40 \rfloor}^{\lceil L/20 \rceil} L^{2\rho}\left(1 - \frac{5}{2L}\right)^L$$

$$\ge \frac{1}{C} L^{2\rho+1}, \tag{4.24}$$

$$c_i^2|z|^{2i} \le Ci^{2\rho}\left(1 - \frac{1}{2L}\right)^{2i} \le Ci^{2\rho}e^{-i/L}$$

$$\le C\max(1, L^{2\rho}) \le CL^{-2\alpha_1}\mathbf{Var}\, P(z) \qquad \forall 0 \le i \le n,$$

where the next to last inequality follows from the boundedness of the function $x \mapsto x^{2\rho}e^{-x}$ on $[0, \infty)$ whenever $\rho \ge 0$ and is trivial when $\rho < 0$. □

Combining Theorem 4.11 and Lemma 4.12, we obtain the following.

PROPOSITION 4.13 (Log-comparability). *There exist constants* $\alpha_0 > 0$ *and* $C > 0$ *such that for every* $\delta \in [\frac{1}{20n}, \frac{1}{C}]$, $1 \le m \le \delta^{-\alpha_0}$, $|z_1|, \dots, |z_m| \in I(\delta) + [-\delta/2, \delta/2]$, *and smooth function* $F : \mathbb{C}^m \to \mathbb{C}$ *with* $\|\nabla^a F\| \le \delta^{-\alpha_0}$, $\forall 0 \le a \le 3$, *we have*

$$\left|\mathbf{E}F\left(\log\left|P(z_1)\right|, \dots, \log\left|P(z_m)\right|\right) - \mathbf{E}F\left(\log\left|\widetilde{P}(z_1)\right|, \dots, \log\left|\widetilde{P}(z_m)\right|\right)\right| \le C\delta^{\alpha_0},$$

and if $\delta \in [\frac{1}{20n}, \frac{1}{C}]$, *we have*

$$\left|\mathbf{E}F\left(\log\left|Q(z_1)\right|, \dots, \log\left|Q(z_m)\right|\right) - \mathbf{E}F\left(\log\left|\widetilde{Q}(z_1)\right|, \dots, \log\left|\widetilde{Q}(z_m)\right|\right)\right| \le C\delta^{\alpha_0}.$$

4.4. *On the tail event* $\mathcal{T}^c$. In this section, we show that if $\mathcal{T}$ is any event such that $\mathbf{P}(\mathcal{T}^c) \le C\gamma(\delta)$ then the contribution from $\mathcal{T}^c$ is negligible. We make use of the powerful result in [27] to deal with the case when the ξ_i's are symmetric. The general case requires some additional tricks in the end.

LEMMA 4.14. *There exists some constant C such that for all $\delta \in [\frac{1}{20n}, \frac{1}{C}]$, one has*

$$\mathbf{E}\left(\left|\prod_{j=1}^{k} X_j^P\right| \mathbf{1}_{\mathcal{T}^c}\right) \le C\delta^{1/22}, \tag{4.25}$$

and when $\delta \in [\frac{1}{10n}, \frac{1}{C}]$, one has

$$\mathbf{E}\left(\left|\prod_{j=1}^{k} X_j^Q\right| \mathbf{1}_{\mathcal{T}^c}\right) \le C\delta^{1/22}. \tag{4.26}$$

PROOF. We will consider two cases.

Case 1. $\frac{\log^2 n}{n} \le \delta \le \frac{1}{C}$. We have $\mathbf{P}(\mathcal{T}^c) \le C\gamma(\delta) = C\delta^A$. By Proposition 4.6, there exists a constant C_1 such that for any $1 \le \lambda \le \frac{n}{L\log^2 L}$,

$$N_{\check{P}}\big(B(\check{z}_j, r_0)\big) = N_P\big(B(z_j, 10^{-5}\delta)\big) \le \lambda\delta^{-c_1},$$

with probability at least $1 - C_1\frac{1}{\lambda^A L^A}$. Hence, $|X_j^P| \le \delta^{-c_1}$ with probability at least $1 - C_1\frac{1}{\lambda^A L^A}$.

For each i such that $i_0 > i \ge 1$, where $2^{i_0-1} \le \frac{n}{L\log^2 L} < 2^{i_0}$, let Ω_i be the set of $\omega \in \mathcal{T}^c$ such that:

(i) $2^{i-1}\delta^{-c_1} < N_{\check{P}}B(\check{z}_j, r_0)$ for some $1 \le j \le k$, and
(ii) $N_{\check{P}}B(\check{z}_j, r_0) \le 2^i\delta^{-c_1}$, $\forall 1 \le j \le k$.

Let $\Omega_0 = \{\omega \in \mathcal{T}^c : N_{\check{P}}B(\check{z}_j, r_0) \le \delta^{-c_1} \forall 1 \le j \le k\}$ and

$$\Omega_{i_0} = \big\{\omega \in \mathcal{T}^c : 2^{i_0-1}\delta^{-c_1} < N_{\check{P}}B(\check{z}_j, r_0) \text{ for some } 1 \le j \le k\big\}.$$

Then $\mathcal{T}^c = \bigcup_{i=0}^{i_0}\Omega_i$, $\mathbf{P}(\Omega_i) \le \frac{C_1\delta^A}{2^{(i-1)A}}$ for all $i \le i_0$ and $|X_j^P| \le 2^i\delta^{-c_1}$ on Ω_i for all $i < i_0$, and $|X_j^P| \le n$ on Ω_{i_0}.

Using the assumption that $\frac{\log^2 n}{n} \le \delta$ and $A \ge k + 2$, we have

$$\begin{aligned}
\mathbf{E}\left(\left|\prod_{j=1}^{k} X_j^P\right| \mathbf{1}_{\mathcal{T}^c}\right) &\le \sum_{i=0}^{i_0-1}\mathbf{E}\left(\left|\prod_{j=1}^{k} X_j^P\right| \mathbf{1}_{\Omega_i}\right) + \mathbf{E}\left(\left|\prod_{j=1}^{k} X_j^P\right| \mathbf{1}_{\Omega_{i_0}}\right) \\
&\le \sum_{i=0}^{\infty}\frac{C_1\delta^A}{2^{(i-1)A}}\big(2^i\delta^{-c_1}\big)^k + \frac{C_1 n^k\delta^A}{2^{(i_0-1)A}} \\
&\le C_1\delta^{A-kc_1}\sum_{i=1}^{\infty}\frac{1}{(2^{A-k})^i} + \frac{C_1 n^k\delta^A}{(n/(2L\log^2 L))^A} \le C\delta^{1/22}.
\end{aligned}$$

Case 2. $\frac{1}{20n} \le \delta \le \frac{\log^2 n}{n}$. Then we have $\mathbf{P}(\mathcal{T}^c) \le C\gamma(\delta) = Cn^{-1/2}$ and $|z_j| \in [1 - \frac{2\log^2 n}{n}, 1 + \frac{1}{n}]$ for all $1 \le j \le k$.

Since ξ_i's satisfy Condition 1(1), there exist positive constants d and q such that $\mathbf{P}(|\xi_i| < d) \le q < 1$. Indeed, if for some $d > 0$, $\mathbf{P}(|\xi_i| < d) > 1 - d$, then by Hölder's inequality,

$$\begin{aligned} 1 &\le \mathbf{E}|\xi_i|^2 = \mathbf{E}|\xi_i|^2 \mathbf{1}_{|\xi_i|<d} + \mathbf{E}|\xi_i|^2 \mathbf{1}_{|\xi_i|\ge d} \\ &\le d^2 + d^{\epsilon/(2+\epsilon)} (\mathbf{E}|\xi_i|^{2+\epsilon})^{2/(2+\epsilon)} \le d^2 + d^{\epsilon/(2+\epsilon)} \tau_2^{2/(2+\epsilon)}. \end{aligned} \tag{4.27}$$

Thus, one can choose d small enough (depending on τ_2 and ϵ), and $q = 1 - d$ to have $\mathbf{P}(|\xi_i| < d) \le q < 1$.

Subcase 2.1. We first consider the case when the random variables ξ_i's are symmetric. In other words, ξ_i and $-\xi_i$ have the same distribution.

Let

$$\mathcal{V} = \{\omega \in \mathcal{T}^c : |\xi_i| \ge d \text{ for some } i \in [N_0, n]\}. \tag{4.28}$$

Since $|X_j^P| \le n$, one has

$$\begin{aligned} \mathbf{E}\left(\left|\prod_{j=1}^k X_j^P\right| \mathbf{1}_{\mathcal{T}^c \setminus \mathcal{V}}\right) &\le n^k \mathbf{P}(|\xi_i| < d, \forall i \in [N_0, n]) \\ &\le n^k q^{n-N_0} \le \frac{1}{20n} \le \delta, \end{aligned} \tag{4.29}$$

when n is sufficiently large. Thus, it suffices to show that

$$\mathbf{E}\left(\left|\prod_{j=1}^k X_j^P\right| \mathbf{1}_{\mathcal{V}}\right) \le C'\delta^{1/22}. \tag{4.30}$$

By Hölder's inequality, we have

$$\mathbf{E}\left(\left|\prod_{j=1}^k X_j^P\right| \mathbf{1}_{\mathcal{V}}\right) \le \prod_{j=1}^k \mathbf{E}(|X_j^P|^k \mathbf{1}_{\mathcal{V}})^{1/k}. \tag{4.31}$$

And so, we reduce the problem to showing that

$$\mathbf{E}|X_j^P|^k \mathbf{1}_{\mathcal{V}} \le C\delta^{1/22} \qquad \forall 1 \le j \le k. \tag{4.32}$$

From (4.6) and the change of variables formula, we obtain

$$\begin{aligned} |X_j^P| &\le CL^2 \int_{B(z_j, 10^{-5}\delta)} \left|\log|P(z)| H_j\left(\frac{z}{10^{-3}\delta}\right)\right| dz \\ &\le CL^2 \int_{B(z_j, 10^{-5}\delta)} |\log|P(z)|| \, dz. \end{aligned} \tag{4.33}$$

And from Hölder's inequality, we have

$$\begin{aligned}\mathbf{E}|X_j^P|^k \mathbf{1}_{\mathcal{V}} &\le CL^{2k}\int_{\mathcal{V}}\Big(\int_{B(z_j,10^{-5}\delta)} |\log|P(z)||\,dz\Big)^k d\mathbf{P}\\ &\le CL^{2k}\int_{\mathcal{V}}\Big(\int_{B(z_j,10^{-5}\delta)} |\log|P(z)||^k\,dz\Big)\\ &\quad\times |B(z_j,10^{-5}\delta)|^{k-1}\,d\mathbf{P},\\ &\le CL^2\int_{\mathcal{V}}\int_{B(z_j,10^{-5}\delta)} |\log|P(z)||^k\,dz\,d\mathbf{P},\\ &\le CL^2\Big(\int_{\mathcal{V}}\int_{B(z_j,10^{-5}\delta)} |\log|P(z)||^{kp}\,dz\,d\mathbf{P}\Big)^{1/p}\\ &\quad\times\Big(\int_{\mathcal{V}}\int_{B(z_j,10^{-5}\delta)} 1\,dz\,d\mathbf{P}\Big)^{1/q},\end{aligned}$$

where p and q are positive constants to be chosen later so that $\frac{1}{p}+\frac{1}{q}=1$. Since $B(z_j,10^{-5}\delta)\subset A(0,1-\frac{3\log^2 n}{n},1+\frac{2}{n}) =: \mathcal{D}$, it follows that

$$\mathbf{E}|X_j^P|^k \mathbf{1}_{\mathcal{V}} \le CL^2\Big(\int_{\mathcal{V}}\int_{\mathcal{D}} |\log|P(z)||^{kp}\,dz\,d\mathbf{P}\Big)^{1/p}\Big(\frac{1}{\sqrt{n}L^2}\Big)^{1/q}. \tag{4.34}$$

For each $N_0\le i\le n$, let $\mathcal{V}_i=\{\omega\in\Omega:|\xi_i|\ge d\}$. By (4.28), $\mathcal{V}\subset\bigcup_{i=N_0}^n \mathcal{V}_i$. Note that this bound is very generous because the measure of $\mathcal{V}_i$ can be very big. Let $I_i=\int_{\mathcal{V}_i}\int_{\mathcal{D}}|\log|P(z)||^{kp}\,dz\,d\mathbf{P}$. Then

$$\int_{\mathcal{V}}\int_{\mathcal{D}} |\log|P(z)||^{kp}\,dz\,d\mathbf{P} \le \sum_{i=N_0}^n I_i =: I. \tag{4.35}$$

Fix $N_0\le i_0\le n$. We will upper bound I_{i_0}.

Let $(\epsilon_m)_{m\in\mathbf{Z}}$ be independent Rademacher random variables (independent of all previous random variables). In [27], Corollary 2.2, Nazarov, Nishry and Sodin showed the following.

THEOREM 4.15. *There exists a constant C_1 such that for any $g(\theta)=\sum_{j\in\mathbf{Z}} a_j\epsilon_j e^{\sqrt{-1}2\pi j\theta}$ with deterministic coefficients a_j's satisfying $\sum_{j\in\mathbf{Z}}|a_j|^2=1$, and any $p_0\ge 1$, one has*

$$\mathbf{E}\int_0^1 |\log|g(\theta)||^{p_0}\,d\theta \le (C_1 p_0)^{6p_0}. \tag{4.36}$$

As a consequence, for any complex numbers $a_0, \dots, a_n$, by Minkowski's inequality for $L^p((\Omega \times [0,1), \mathbf{P} \times m))$, we have

$$\begin{aligned}
&\mathbf{E}\int_0^1 \left|\log\left|\sum_{j=0}^n a_j \epsilon_j e^{\sqrt{-1}2\pi j\theta}\right|\right|^{p_0} d\theta \\
&\qquad \le \left((C_1 p_0)^6 + \frac{1}{2}\left|\log \sum_{j=0}^n |a_j|^2\right|\right)^{p_0} \qquad (4.37)\\
&\qquad \le 2^{p_0}(C_1 p_0)^{6p_0} + \left|\log \sum_{j=0}^n |a_j|^2\right|^{p_0}.
\end{aligned}$$

Let $\hat{\xi}_k = \epsilon_k \xi_k$. Since ξ_k is symmetric, $\hat{\xi}_k$ has the same distribution as ξ_k. And so, the random variables $\hat{\xi}_0, \dots, \hat{\xi}_n$ have the same joint distribution as $\xi_0, \dots, \xi_n$. Thus, from the definition of I_{i_0}, we have

$$\begin{aligned}
I_{i_0} &= \int_{|\hat{\xi}_{i_0}| \ge d} \int_{\mathcal{D}} \left|\log\left|\sum_{j=0}^n c_j \hat{\xi}_j z^j\right|\right|^{kp} dz\, d\mathbf{P} \\
&= 2\pi \int_{1-\frac{3\log^2 n}{n}}^{1+2/n} r \int_{|\xi_{i_0}| \ge d} \int_0^1 \left|\log\left|\sum_{j=0}^n c_j \xi_j \epsilon_j r^j e^{\sqrt{-1}2\pi j\theta}\right|\right|^{kp} d\theta\, d\mathbf{P}\, dr.
\end{aligned}$$

Conditioning on the event $|\xi_{i_0}| \ge d$ and fixing the ξ_i's, from (4.37), we obtain

$$\begin{aligned}
&\mathbf{E}_{\epsilon_0,\dots,\epsilon_n} \int_0^1 \left|\log\left|\sum_{j=0}^n c_j \xi_j \epsilon_j r^j e^{\sqrt{-1}2\pi j\theta}\right|\right|^{kp} d\theta \\
&\qquad \le (2C_1 kp)^{6kp} + \left|\log \sum_{j=0}^n |c_j \xi_j r^j|^2\right|^{kp}.
\end{aligned}$$

Undoing the conditioning, we get

$$\begin{aligned}
I_{i_0} &\le 2\pi \int_{1-\frac{3\log^2 n}{n}}^{1+2/n} r\left((Cp)^{6kp} + \int_{|\xi_{i_0}| \ge d} \left|\log \sum_{j=0}^n |c_j j \xi_j r^j|^2\right|^{kp} d\mathbf{P}\right) dr \qquad (4.38)\\
&\le C + C\int_{1-\frac{3\log^2 n}{n}}^{1+2/n} r \int_{|\xi_{i_0}| \ge d} \left|\log \sum_{j=0}^n |c_j r^j \xi_j|^2\right|^{kp} d\mathbf{P}\, dr.
\end{aligned}$$

By (2.5), for every $N_0 \le j \le n$ and $1 - \frac{3\log^2 n}{n} \le r \le 1 + \frac{2}{n}$, we have $\frac{1}{Cn^{2|\rho|}} r^{2j} \le c_j^2 r^{2j} \le Cn^{2|\rho|}$. And hence, on the event $|\xi_{i_0}| \ge d$,

$$\frac{1}{Cn^{2|\rho|}} r^{2i_0} d^2 \le \sum_{j=0}^n |c_j r^j \xi_j|^2 \le Cn^{2|\rho|+1} \sum_{j=0}^n \xi_j^2.$$

Hence,

$$\left|\log \sum_{j=0}^{n} |c_j r^j \xi_j|^2\right| \le \max\left\{\left|\log\left(\frac{1}{Cn^{2|\rho|}} r^{2i_0} d^2\right)\right|, \left|\log\left(Cn^{2|\rho|+1} \sum_{j=0}^{n} \xi_j^2\right)\right|\right\}.$$

And so,

$$\begin{aligned}
&\mathbf{E}\mathbf{1}_{|\xi_{i_0}|\ge d}\left|\log \sum_{j=0}^{n} |c_j r^j \xi_j|^2\right|^{kp} \\
&\quad\le C\mathbf{E}\mathbf{1}_{|\xi_{i_0}|\ge d}\left|\log \sum_{j=0}^{n} \xi_j^2\right|^{kp} + C\log^{kp} n + C\big|\log|r^{2i_0}d^2|\big|^{kp} \\
&\quad= C\mathbf{E}\mathbf{1}_{|\xi_{i_0}|\ge d, \sum_{j=0}^{n} \xi_j^2 \ge 1}\left|\log \sum_{j=0}^{n} \xi_j^2\right|^{kp} \\
&\qquad+ C\mathbf{E}\mathbf{1}_{|\xi_{i_0}|\ge d, \sum_{j=0}^{n} \xi_j^2 < 1}\left|\log \sum_{j=0}^{n} \xi_j^2\right|^{kp} \\
&\qquad+ C\log^{kp} n + C\big|\log|r^{2i_0}d^2|\big|^{kp} \\
&\quad\le C\log^{kp} n + C|\log d^2|^{kp} + C\big|\log|r^{2i_0}d^2|\big|^{kp}.
\end{aligned}$$

Notice that under Condition 1, the coefficients c_i of P, thanks to their polynomial growth, only contribute the term $\log^{kp} n$ in the above estimate. The same applies for Q because $\frac{C}{n^{|\rho|}} \le |\frac{d_i}{d_0}| \le Cn^{|\rho|}$.

Plugging in (4.38) gives that for all $N_0 \le i_0 \le n$, one has

$$I_{i_0} \le C\log^{kp} n + C\int_{1-\frac{3\log^2 n}{n}}^{1+2/n} r|\log r^{2i_0} d^2|^{kp}\, dr \le C\log^{kp} n, \tag{4.39}$$

and so

$$\begin{aligned}
\int_{\exists i\in[N_0,n]:|\xi_i|\ge d}\int_{\mathcal{D}}\left|\log\left|\sum_{j=0}^{n} c_j\xi_j z^j\right|\right|^{kp} dz \le I = \sum_{i=N_0}^{n} I_i& \\
\le C_2 n\log^{kp} n.&
\end{aligned} \tag{4.40}$$

We now use a rude bound which will be convenient for the next subcase:

$$\int_{\exists i\in[N_0,n]:|\xi_i|\ge d}\int_{\mathcal{D}}\left|\log\left|\sum_{j=0}^{n} c_j\xi_j z^j\right|\right|^{kp} dz \le I \le C_2 n^{4/3}\log^{kp} n. \tag{4.41}$$

Combining this with (4.34) and (4.35), we obtain

$$\begin{aligned}
&\mathbf{E}|X_j^P|^k \mathbf{1}_{\mathcal{V}} \\
&\quad \le CL^2 I^{1/p}\left(C\frac{1}{\sqrt{n}L^2}\right)^{1/q} \le CL^2 n^{4/3p}\log^k n\left(\frac{1}{\sqrt{n}L^2}\right)^{1/q} \\
&\quad \le CL^2 L^{8/3p}(\log L^2)^k\left(\frac{1}{L^{5/2}}\right)^{1/q} \\
&\qquad \left(\text{since } \sqrt{n} \le \frac{n}{\log^2 n} \le L \le 20n\right) \\
&\quad = C\frac{\log^k L}{L^{\frac{5}{2q}-2-\frac{8}{3p}}} = C\frac{\log^k L}{L^{\frac{1}{2}-\frac{31}{6p}}} \le \frac{C}{L^{1/22}} \\
&\qquad \left(\text{by choosing } p = 12, q = \frac{12}{11}\right).
\end{aligned} \tag{4.42}$$

This together with (4.31) complete the proof of (4.30).

Subcase 2.2. Now let us consider the case when the ξ_i's are not symmetric. The trick is to reduce to the symmetric case. For clarity, we write $P_\xi(z) = \sum_{l=0}^n c_l \xi_l z^l$.

Recall that d and q are constants such that $\mathbf{P}(|\xi_i| < d) \le q < 1$. Let $\xi_0', \dots, \xi_n'$ be independent copies of $\xi_0, \dots, \xi_n$ correspondingly. For this subcase, instead of (4.28), we set

$$\mathcal{V} = \{\omega \in \mathcal{T}^c : |\xi_i| \ge d, |\xi_i'| \ge d \text{ for some } i \in [N_0, n]\}. \tag{4.43}$$

Correspondingly,

$$\mathcal{V}_i = \{\omega \in \Omega : |\xi_i| \ge d, |\xi_i'| \ge d\}. \tag{4.44}$$

Let $\bar{\xi}_l = \frac{\xi_l - \xi_l'}{\sqrt{2}}$. Then the $\bar{\xi}_l$'s are symmetric and satisfy Condition 1(1) (with a different τ_2). Let $\bar{d} < 1$ and $\bar{q}$ be positive constants such that $\mathbf{P}(|\bar{\xi}_1| < \bar{d}) \le \bar{q} < 1$ for all i.

In the following, we will show that

$$I_i = \int_{\mathcal{V}_i}\int_{\mathcal{D}} |\log|P_\xi(z)||^{kp}\, dz\, d\mathbf{P} \le 3n^{1/3}\log^{10kp} n =: 3K_0 \tag{4.45}$$

for all $N_0 \le i \le n$, where $p = 12$ and then, one can use the same argument as in the symmetric case to complete the proof.

Let

$$j_0 = \left\lceil \frac{1}{\bar{q}^{(n+1)/(4kp+8)}} \right\rceil. \tag{4.46}$$

We will first show that

$$\int_{\mathcal{V}_i}\int_{\mathcal{D}} |\log|P_\xi(z)||^{kp}\mathbf{1}_{B_\xi}\, dz\, d\mathbf{P} \le K_0 \qquad \text{for all } N_0 \le i \le n, \tag{4.47}$$

where $B_\xi = \{(\omega, z) \in \mathcal{V}_i \times \mathcal{D} : |\log|P_\xi(z)|| \geq j_0\}$. Indeed, by Hölder's inequality, one has

$$\begin{aligned}
&\int_{\mathcal{V}_i}\int_{\mathcal{D}} \big|\log\big|P_\xi(z)\big|\big|^{kp} \mathbf{1}_{B_\xi}\, dz\, d\mathbf{P} \\
&\quad \leq \biggl(\int_{\mathcal{V}_i}\int_{\mathcal{D}} \big|\log\big|P_\xi(z)\big|\big|^{2kp}\, dz\, d\mathbf{P}\biggr)^{1/2} \biggl(\int_{\mathcal{V}_i}\int_{\mathcal{D}} \mathbf{1}_{B_\xi}\, dz\, d\mathbf{P}\biggr)^{1/2}.
\end{aligned} \tag{4.48}$$

To bound the first integral on the right, let $\epsilon_0', \dots, \epsilon_n'$ be independent Rademacher variables defined on $(\{\pm 1\}^{n+1}, \nu)$ where ν is the uniform probability measure on $\{\pm 1\}^{n+1}$. Let $(\hat{\Omega}, \hat{\mu}) = (\Omega \times \{\pm 1\}^{n+1}, \mathbf{P} \times \nu)$, and define the random variables $\hat{\xi}_i(\omega_1, \omega_2) = \xi_i(\omega_1)\epsilon_i'(\omega_2)$ for all $\omega_1 \in \Omega$ and $\omega_2 \in \{\pm 1\}^{n+1}$.

Observe that $\hat{\xi}_i'$ is symmetric and equal to ξ_i when $\epsilon_i' = 1$. Let $s > 1$ be any constant such that $2^{1/s} \leq \frac{1}{\bar{q}^{kp/(4kp+8)}}$. We have

$$\begin{aligned}
&\int_{\mathcal{V}_i}\int_{\mathcal{D}} \big|\log\big|P_\xi(z)\big|\big|^{2kp}\, dz\, d\mathbf{P} \\
&\quad = 2^{n+1} \int_{\{1\}^{n+1}} \int_{\mathcal{V}_i}\int_{\mathcal{D}} \big|\log\big|P_\xi(z)\big|\big|^{2kp}\, dz\, d\mathbf{P}\, d\nu \\
&\quad = 2^{n+1} \int_{\mathcal{V}_i \times \{1\}^{n+1}} \int_{\mathcal{D}} \big|\log\big|P_{\hat{\xi}}(z)\big|\big|^{2kp}\, dz\, d\hat{\mu} \qquad \text{by Fubini's theorem} \\
&\quad \leq 2^{n+1} \big[\mathcal{D}\hat{\mu}\big(\mathcal{V}_i \times \{1\}^{n+1}\big)\big]^{1-1/s} m \biggl(\int_{\mathcal{V}_i \times \{1\}^{n+1}} \int_{\mathcal{D}} \big|\log\big|P_{\hat{\xi}}(z)\big|\big|^{2kps}\, dz\, d\hat{\mu}\biggr)^{1/s} \\
&\qquad \text{(by Hölder's inequality)} \\
&\quad \leq 2^{(n+1)/s} \biggl(\int_{\mathcal{V}_i \times \{1\}^{n+1}} \int_{\mathcal{D}} \big|\log\big|P_{\hat{\xi}}(z)\big|\big|^{2kps}\, dz\, d\hat{\mu}\biggr)^{1/s} \\
&\qquad \big(\text{because } \hat{\mu}(\mathcal{V}_i \times \{1\}^{n+1}) \leq 2^{-n-1}\big) \\
&\quad \leq C 2^{(n+1)/s} \log^{2kp} n \qquad \text{by (4.39) for } \hat{\xi}_i.
\end{aligned}$$

A bound for the second integral on the right of (4.48) can also be derived from the above bound.

$$\int_{\mathcal{V}_i}\int_{\mathcal{D}} \mathbf{1}_{B_\xi}\, dz\, d\mathbf{P} \leq \frac{1}{j_0^{2kp}} \int_{\mathcal{V}_i}\int_{\mathcal{D}} \big|\log\big|P_\xi(z)\big|\big|^{2kp}\, dz\, d\mathbf{P} \leq C \frac{2^{n/s}}{j_0^{2kp}} \log^{2kp} n.$$

Plugging into (4.48) gives

$$\begin{aligned}
\int_{\mathcal{V}_i}\int_{\mathcal{D}} \big|\log\big|P_\xi(z)\big|\big|^{kp} \mathbf{1}_{B_\xi}\, dz\, d\mathbf{P} &\leq C \frac{2^{(n+1)/s}}{j_0^{kp}} \log^{2kp} n \\
&\leq C \log^{2kp} n < K_0,
\end{aligned} \tag{4.49}$$

where the next to last inequality follows directly from the way we set s and j_0. This proves (4.47).

Now assume to the contrary that (4.45) failed, that is, $I_i > 3K_0$ for some i. Thanks to (4.47), one then has

$$\int_{\mathcal{V}_i}\int_{\mathcal{D}} \left|\log\left|P_\xi(z)\right|\right|^{kp} \mathbf{1}_{|\log|P_\xi(z)||<j_0}\, dz\, d\mathbf{P} > 2K_0. \tag{4.50}$$

For each $z \in \mathcal{D}$ and $1 \le j \le j_0$, set $\mu_z(j) = \mathbf{P}(\mathcal{V}_i \cap (j-1 \le |\log|P_\xi(z)|| < j))$. Since

$$\begin{aligned} 2K_0 < I_i &\le \sum_{j=1}^{j_0} \int_{\mathcal{D}} j^{kp} \mu_\xi(j)\, d\mathbf{P} \\ &\le \sum_{j=1}^{j_0} j^{kp} m\left(\mathcal{D} \cap \left\{\omega : \mu_\xi(j) \ge \frac{1}{j^{kp+2}}\right\}\right) + \sum_{j=1}^{\infty} \frac{1}{j^2}, \end{aligned}$$

and $\sum_{j=1}^{\infty} \frac{1}{j^2} \le 2$, there exists a number $j \le j_0$ such that

$$1 \ge m(\mathcal{D}) \ge m(\mathcal{D}_0) \ge \frac{K_0}{2j^{kp+2}}, \tag{4.51}$$

where $\mathcal{D}_0 = \{z \in \mathcal{D} : \mu_\xi(j) \ge \frac{1}{j^{kp+2}}\}$. Since $j^{kp+2} \ge \frac{K_0}{2} \ge \frac{\log^{10kp} n}{2}$, we have $j \ge \log^5 n$. Now, by Markov's inequality and Condition 1, for any $z \in \mathcal{D}$,

$$\begin{aligned} \mathbf{P}\left(\log\left|P(z)\right| \ge j-1\right) &\le \frac{\mathbf{E}|P(z)|^2}{e^{2j-2}} \le \frac{1}{e^{2j-2}}\left(\sum_{i=0}^{n} |c_i||z|^i \left(\mathbf{E}|\xi_i|^2\right)^{1/2}\right)^2 \\ &\le \frac{n^{2\rho+2}}{e^{2j-2}} \le \frac{1}{e^j} \le \frac{1}{2j^{kp+2}}. \end{aligned}$$

Thus, $p_z := \mathbf{P}(\mathcal{V}_i \cap (-j < \log|P_\xi(z)| \le -j+1)) \ge \frac{1}{2j^{kp+2}}$ for every $z \in \mathcal{D}_0$.

On the set $\mathcal{D}_0$,

$$\begin{aligned} &\mathbf{P}\left(\omega \in \mathcal{V}_i : -j < \log\left|P_\xi(z)\right|, \log\left|P_{\xi'}(z)\right| \le -j+1 \text{ and } \exists i' \in [N_0, n] : |\bar{\xi}_{i'}| \ge \bar{d}\right) \\ &\quad \ge \mathbf{P}\left(\omega \in \mathcal{V}_i : -j < \log\left|P_\xi(z)\right|, \log\left|P_{\xi'}(z)\right| \le -j+1\right) - \mathbf{P}\left(|\bar{\xi}_{i'}| < \bar{d}, \forall i'\right) \\ &\quad \ge p_z^2 - \bar{q}^{n+1}. \end{aligned}$$

From the definition (4.46) of j_0, we have

$$\bar{q}^{n+1} \le \frac{1}{(j_0-1)^{4kp+8}} \le \frac{1}{2} p_z^2, \tag{4.52}$$

and thus, $p_z^2 - \bar{q}^{n+1} \geq \frac{1}{2} p_z^2$. Therefore, on $\mathcal{D}_0$ we have

$$\mathbf{P} \times m\big((\omega, z) \in \bar{\mathcal{U}} \times \mathcal{D}_0 : -j < \log\big|P_\xi(z)\big|, \log\big|P_{\xi'}(z)\big| \leq -j+1\big) \geq \int_{\mathcal{D}_0} \frac{1}{2} p_z^2 \, dz \geq \frac{1}{2}\bigg(\frac{K_0}{16 j^{2kp+4}}\bigg)^3,$$

where $\bar{\mathcal{U}} = \{\omega : \exists i' \in [N_0, n] : |\bar{\xi}_{i'}| \geq \bar{d}\}$.

Note that when $-j < \log|P_\xi(z)|$, $\log|P_{\xi'}(z)| \leq -j+1$, we have $|P_{\bar{\xi}}(z)| \leq \sqrt{2} e^{-j+1}$, so $\log|P_{\bar{\xi}}(z)| \leq -\frac{j}{2}$. This implies

$$\int_{\bar{\mathcal{U}}} \int_{\mathcal{D}} \big|\log\big|P_{\bar{\xi}}(z)\big|\big|^{6kp+12} \geq \frac{1}{2}\bigg(\frac{j}{2}\bigg)^{6kp+12}\bigg(\frac{K_0}{16 j^{2kp+4}}\bigg)^3 = \frac{K_0^3}{2^{6kp+25}} = \frac{n \log^{30kp} n}{2^{6kp+25}}. \tag{4.53}$$

Now since $\bar{\xi}_i$'s are symmetric and satisfy Condition 1(1), (4.40) holds for $\bar{\xi}_i$ with $\bar{d}$ in place of d and $6kp+12$ in place of kp and gives

$$\int_{\bar{\mathcal{U}}} \int_{\mathcal{D}} \big|\log\big|P_{\bar{\xi}}(z)\big|\big|^{6kp+12} \leq C n \log^{6kp+12} n. \tag{4.54}$$

Now as $p = 12$, the bounds (4.53) and (4.54) provide a contradiction which then completes the proof of Lemma 4.14. □

4.5. *Finishing*. Finally, we will complete the proof of Theorem 2.3.

Let φ_0 be a smooth function on $\mathbb{C}^k$ such that $\varphi_0(z_1, \ldots, z_k) = z_1 \cdots z_k$ on $B(0, \delta^{-c_1})^k$, $= 0$ outside of $B(0, 2\delta^{-c_1})^k$, $|\varphi_0(z_1, \ldots, z_k)| \leq |z_1| \cdots |z_k|$ for all $(z_1, \ldots, z_k) \in \mathbb{C}^k$ and $|\nabla^a \varphi_0(\omega)| \leq C\delta^{-kc_1}$ for all $0 \leq a \leq 3$. For example, $\varphi_0(z_1, \ldots, z_k) = \prod_{i=1}^k z_i \phi(\frac{|z_i|}{\delta^{-c_1}})$ for some smooth function ϕ such that ϕ is a smooth function such that $\operatorname{supp}(\phi) \subset [-2, 2]$, $0 \leq \phi \leq 1$, and $\phi = 1$ on $[-1, 1]$.

Since $X_j^P \leq \delta^{-c_1}$ on $\mathcal{T}$, we have

$$\begin{aligned}
&\mathbf{E}_\xi \left|\prod_{j=1}^k X_j^P - \varphi_0\big(X_1^P, \ldots, X_k^P\big)\right| \\
&\quad = \mathbf{E}_\xi \left|\prod_{j=1}^k X_j^P - \varphi_0\big(X_1^P, \ldots, X_k^P\big)\right| \mathbf{1}_{\mathcal{T}^c} \\
&\quad \leq 2\mathbf{E}_\xi \left(\left|\prod_{j=1}^k X_j^P\right| \mathbf{1}_{\mathcal{T}^c}\right) \leq C'\delta^{1/22} \qquad \text{by (4.25)},
\end{aligned}$$

where by $\mathbf{E}_\xi$, we mean the expectation with respect to the random variables $\xi_0, \ldots, \xi_n$.

From Proposition 4.7 and (4.20), we deduce that on the product space generated by the random variables $\xi_0, \dots, \xi_n$ and the random points $\check{w}_{j,i}$, the bound (4.20) holds with probability at least $1 - \gamma_0 - C\gamma(\delta)$. Thus,

$$\mathbf{E}_{\xi,\check{w}}\left|\varphi_0(X_1^P, \dots, X_k^P) - \varphi_0\left(\frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_1^P(\check{w}_{1,i}), \dots, \frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_k^P(\check{w}_{k,i})\right)\right|$$

$$\leq C\left(\delta^{-kc_1}\frac{L^{4c_1}}{\sqrt{m_0\gamma_0}} + \delta^{-kc_1}(\gamma(\delta) + k\gamma_0)\right),$$

where $\mathbf{E}_{\xi,\check{w}}$ is the expectation on the product space. The first term bounds the contribution of the good event when (4.20) holds, and follows from the bound on the first derivative of φ_0. The second term bounds the contribution of the bad event when (4.20) fails and follows from the bound on the infinity norm of φ_0.

Note that $\gamma(\delta) \leq 10\delta^{1/2}$ for all $\frac{1}{20n} \leq \delta \leq \frac{1}{C}$.

Let c be any constant such that $0 < c \leq \min\{\frac{1}{22}, \frac{\alpha_0}{2(3k+11)}, \frac{1}{2(k+1)}\}$ where α_0 is the constant in Proposition 4.13. Let $c_1 = c$, $m_0 = \lfloor \delta^{-(3k+11)c} \rfloor$, and $\gamma_0 = \delta^{(k+1)c}$, then the above error term is $C\delta^c$, and so

$$\mathbf{E}_{\xi,\check{w}}\left|\prod_{j=1}^{k} X_j^P - \varphi_0\left(\frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_1^P(\check{w}_{1,i}), \dots, \frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_k^P(\check{w}_{k,i})\right)\right| \leq C\delta^c.$$

Now, applying Proposition 4.13 by first conditioning on the points $\check{w}_{j,i}$, we obtain

$$\left|\mathbf{E}_{\xi,\check{w}}\varphi_0\left(\frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_1^P(\check{w}_{1,i}), \dots, \frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_k^P(\check{w}_{k,i})\right)\right.$$

$$\left. - \mathbf{E}_{\tilde{\xi},\check{w}}\varphi_0\left(\frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_1^{\tilde{P}}(\check{w}_{1,i}), \dots, \frac{\pi r_0^2}{m_0}\sum_{i=1}^{m_0} K_k^{\tilde{P}}(\check{w}_{k,i})\right)\right| \leq C\delta^c.$$

This completes the proof of Theorem 2.3.

5. Proof of real local universality. In this section, we will prove Theorem 2.4. As before, we can assume without loss of generality that $\tilde{\xi}_i$ has Gaussian distribution for all i.

Let $r_0 = 10^{-2}/2$. Below, we prove (2.10); the same proof works for Q in place of P unless otherwise noted. As before, we reduce the problem to showing (2.10) for functions G of the form

$$G(y_1, \dots, y_k, w_1, \dots, w_l) = F_1(y_1)\cdots F_k(y_k)G_1(w_1)\cdots G_l(w_l),$$

where $F_i : \mathbb{R} \to \mathbb{C}$ and $G_j : \mathbb{C} \to \mathbb{C}$ are smooth functions supported on $[-r_0, r_0]$ and $B(0, r_0)$, respectively, such that

$$|\nabla^a F_i(x)|, |\nabla^a G_j(z)| \leq 1$$

for all $1 \leq i \leq k$, $1 \leq j \leq l$, $x \in \mathbb{R}$, $z \in \mathbb{C}$, and $0 \leq a \leq 3$.

Then, by the inclusion-exclusion argument and the symmetry of zeros of P about the x-axis, we can further reduce the problem to showing that

$$\left|\mathbf{E}\left(\prod_{j=1}^{k} X^{P}_{\check{x}_i,F_i,\mathbb{R}}\right)\left(\prod_{j=1}^{l} X^{P}_{\check{z}_j,G_j,\mathbb{C}_+}\right) - \mathbf{E}\left(\prod_{j=1}^{k} X^{\widetilde{P}}_{\check{x}_i,F_i,\mathbb{R}}\right)\left(\prod_{j=1}^{l} X^{\widetilde{P}}_{\check{z}_j,G_j,\mathbb{C}_+}\right)\right| \le C\delta^{c}, \tag{5.1}$$

where $X^{P}_{\check{x}_i,F_i,\mathbb{R}} = \sum_{\zeta_j^{\check{P}}\in\mathbb{R}} F_i(\zeta_j^{\check{P}} - \check{x}_i)$ and $X^{P}_{\check{z}_j,G_j,\mathbb{C}_+} = \sum_{\zeta_i^{\check{P}}\in\mathbb{C}_+} G_j(\zeta_i^{\check{P}} - \check{z}_j)$.

Since the proof of Theorem 2.3 [and in particular, (4.5)] hardly changes if we replace $I(\delta)$ by $I(\delta) + (-10^{-6}\delta, 10^{-6}\delta)$, we conclude that there exists a positive constant c for which

$$\left|\mathbf{E}\left(\prod_{j=1}^{m} X^{P}_{\check{w}_j,H_j}\right) - \mathbf{E}\left(\prod_{j=1}^{m} X^{P}_{\check{w}_j,H_j}\right)\right| \le C\delta^{c}, \tag{5.2}$$

where $1 \le m \le k + l$, $|w_j| \in I(\delta) + (-10^{-4}\delta, 10^{-4}\delta)$, $H_j : \mathbb{C} \to \mathbb{C}$ is a smooth function supported in $B(0, 2r_0)$ and $|\nabla^a H_j| \le 1$, $\forall 0 \le a \le 3$, and $X_{\check{w}_j,H_j} = \sum_{i=1}^{n} H_j(\zeta_i^{\check{P}} - \check{w}_j)$. For the rest of the proof, we will write, for example, $X_{\check{w}_j,H_j}$ when it can be either $X^{P}_{\check{w}_j,H_j}$ or $X^{\widetilde{P}}_{\check{w}_j,H_j}$.

We shall reduce (5.1) to (5.2) by first showing that the number of complex zeros near the real axis is small with high probability. This is the key lemma for this proof. We make use of a more classical tool, the Rouché's theorem, together with some elegant arguments in [13] and [33].

LEMMA 5.1. *Let c be as in (5.2). Let $c_2 = \min\{\frac{c}{100}, \frac{c}{3k+3l+1}, \frac{\rho+1/2}{4}\}$ and $\gamma = \delta^{c_2}$. There exists a constant C such that for all $\frac{1}{20n} \le \delta \le \frac{1}{C}$, one has*

$$\mathbf{P}\big(N_{\check{P}} B(\check{x}, \gamma) \ge 2\big) \le C\gamma^{3/2},$$

for all $x \in \mathbb{R}$ with $|x| \in I(\delta) + (-10^{-4}\delta, 10^{-4}\delta)$.

When $\delta \ge \frac{1}{10n}$, the same statement holds for Q in place of P.

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

PROOF. We will prove the Lemma for P. The same arguments also work for Q unless otherwise noted. The strategy is using Theorem 2.3 to reduce to Gaussian case. Let H be a nonnegative smooth function supported on $B(0, 2\gamma)$, which equals 1 on $B(0, \gamma)$ and is at most 1 everywhere else, and $|\nabla^a H| \le C\gamma^{-a}$ for all $0 \le a \le 8$. In particular, one can take $H(z) = \phi(\frac{z}{\gamma})$ where ϕ is any smooth function supported in $B(0, 2)$ and equals 1 on $B(0, 1)$.

By Theorem 2.3, we have

$$\begin{aligned}
&\mathbf{P}\big(N_{\check{P}} B(\check{x}, \gamma) \geq 2\big) \\
&\quad \leq \mathbf{E} \sum_{i \neq j} H(\check{\zeta}_i - \check{x}) H(\check{\zeta}_j - \check{x}) \\
&\quad \leq \mathbf{E} \sum_{i \neq j} H(\check{\check{\zeta}}_i - \check{x}) H(\check{\check{\zeta}}_j - \check{x}) + C\delta^c \gamma^{-8} \\
&\quad \leq \mathbf{E} \sum_{i \neq j} H(\check{\check{\zeta}}_i - \check{x}) H(\check{\check{\zeta}}_j - \check{x}) \mathbf{1}_{k \geq \delta^{-c_3}} + \mathbf{E} k(k-1) \mathbf{1}_{k < \delta^{-c_3}} + C\gamma^{3/2} \\
&\quad \leq \mathbf{E} \sum_{i \neq j} H(\check{\check{\zeta}}_i - \check{x}) H(\check{\check{\zeta}}_j - \check{x}) \mathbf{1}_{k \geq \delta^{-c_3}} + \delta^{-2c_3} \mathbf{P}(k \geq 2) + C\gamma^{3/2},
\end{aligned}$$

where $k = N_{\check{\tilde{P}}} B(\check{x}, 2\gamma) = N_{\tilde{P}} B(x, 2.10^{-3}\gamma\delta) =: N_{\tilde{P}} B(x, \eta)$, and $c_3 = c_2/10$. By Proposition 4.6, $k \leq \delta^{-c_3}$ with probability at least $1 - C\gamma(\delta)$.

Using the result from Section 4.4, we have

$$\begin{aligned}
\mathbf{E} \sum_{i \neq j} H(\check{\check{\zeta}}_i - \check{x}) H(\check{\check{\zeta}}_j - \check{x}) \mathbf{1}_{k \geq \delta^{-c_3}} &\leq \mathbf{E}\left(\left(\sum_{i=1}^{n} H(\check{\check{\zeta}}_i - \check{x}) \right)^2 \mathbf{1}_{k \geq \delta^{-c_3}} \right) \\
&\leq C\delta^{1/22} \gamma^{-8} \leq C\gamma^{3/2}.
\end{aligned}$$

Thus, it remains to show that $\mathbf{P}(k \geq 2) = \mathbf{P}(N_{\tilde{P}} B(x, \eta) \geq 2) \leq C\delta^{2c_3}\gamma^{3/2}$. Having reduced the task to the Gaussian case, we will adapt the proofs of similar results in [13] and [32] to show it.

Consider $g(z) = \tilde{P}(x) + \tilde{P}'(x)(z - x)$ and put $v_z = (c_i z^i)_{i=0}^n$. Let $p(z) = \tilde{P}(z) - g(z)$. Notice that for this Gaussian case, $\mathbf{P}(\tilde{P}(x) = 0) = 0$ when n is sufficiently large. Since g is linear, it has at most one zero in $B(x, \eta)$, and hence, when $k \geq 2$, $\tilde{P}$ has more zeros than g in that ball. If $|g(z)| > |p(z)|$ for all $z \in \partial B(x, \eta)$, then by Rouché's theorem, $\tilde{P}$ and g have the same number of zeros. Thus, for all $t > 0$, we have

$$\mathbf{P}(k \geq 2) \leq \mathbf{P}\Big(\min_{z \in \partial B(x,\eta)} |g(z)| \leq \max_{z \in \partial B(x,\eta)} |p(z)| \Big).$$

Let $A_1 = \{\omega : \min_{z \in \partial B(x,\eta)} |g(z)| \leq \max_{z \in \partial B(x,\eta)} |p(z)|\}$. We will show that $\mathbf{P}(A_1) \leq C\delta^{2c_3}\gamma^{3/2}$.

We have $p(z) = (\tilde{\xi}_i)_{i=0}^n (v_z - v_x - v'(x)(z - x))$ and

$$\big|\big(v_z - v_x - v'(x)(z - x)\big)_i\big| \leq \sup_{0 \leq \theta \leq 1} \frac{1}{2} |c_i| |z - x|^2 i(i-1) |x + \theta z|^{i-2}. \tag{5.3}$$

If $\delta \geq \frac{1}{10n}$, then for all $z \in \partial B(x,\eta)$ and $\theta \in [0,1]$, $|x+\theta z| \leq 1-\frac{\delta}{2}$, and so by Condition 1,

$$\mathbf{Var}\big(p(z)\big) = \big|v_z - v_x - v'(x)(z-x)\big|^2 \leq \sum_{i=0}^{n} \eta^4 i^4 c_i^2 \left(1-\frac{\delta}{2}\right)^{2i-4} \leq CL^{2\rho+1-4c_2}.$$

Similarly, if $\frac{1}{20n} \leq \delta \leq \frac{1}{10n}$, then for all $z \in \partial B(x,\eta)$ and $\theta \in [0,1]$, $|x+\theta z| \leq 1+\frac{3}{n}$, and so

$$\mathbf{Var}\big(p(z)\big) \leq \sum_{i=0}^{n} \eta^4 i^4 c_i^2 \left(1+\frac{3}{n}\right)^{2i-4} \leq C\sum_{i=0}^{n} \eta^4 n^{4+2\rho}\left(1+\frac{3}{n}\right)^{2n} \leq CL^{2\rho+1-4c_2}.$$

Thus, in any case,

$$\mathbf{Var}\big(p(z)\big) \leq CL^{2\rho+1-4c_2}. \tag{5.4}$$

[When proving the Lemma for $\widetilde{Q}$, there are two cases: if $\rho \geq 0$ then observe from Condition 1 that $|\frac{d_i}{d_0}| \leq C = Ci^0$ for all i, and so, by the same argument as above, for the function $p(z) = \widetilde{Q}(z) - \widetilde{Q}(x) - (z-x)\widetilde{Q}'(x)$, one has $\mathbf{Var}(p(z)) \leq CL^{(2)(0)+1-4c_2} = CL^{1-4c_2}$ which is similar to the case $\rho = 0$ for P. Now, if $-\frac{1}{2} < \rho < 0$ we similarly have

$$\mathbf{Var}\big[p(z)\big] \leq C\eta^4 \sum_{0\leq i\leq n/2} i^4 e^{-\delta i} + C\eta^4 \sum_{n/2<i\leq n} i^4 \frac{(n-i)^{2\rho}}{n^{2\rho}} e^{-\delta i} \leq CL^{1-4c_2}$$

which again is similar to the case $\rho = 0$ for P. We note that in all computation it is very important that $2\rho+1 > 0$ to ensure that the harmonic sum $\sum_{j\leq M} j^{2\rho}$ is dominated by $M^{2\rho+1}$.]

We use the above estimate to prove that for every $t > 0$,

$$\mathbf{P}\Big(\max_{z\in\partial B(x,\eta)} \big|p(z) - \mathbf{E}p(z)\big| \geq t\Big) \leq Ce^{-t^2/(CL^{2\rho+1-4c_2})}. \tag{5.5}$$

Indeed, let $\bar{p}(z) = p(z) - \mathbf{E}p(z)$, then for every $z \in \partial B(x,\eta)$, by Cauchy's integral formula,

$$\big|\bar{p}(z)\big| \leq \int_0^{2\pi} \frac{|\bar{p}(x+2\eta e^{\sqrt{-1}\theta})|}{|z-x-2\eta e^{\sqrt{-1}\theta}|} 2\eta \frac{d\theta}{2\pi} \leq \sqrt{CL^{2\rho+1-4c_2}} \int_0^{2\pi} \frac{|\bar{p}(x+2\eta e^{\sqrt{-1}\theta})|}{\sqrt{\mathbf{Var}(\bar{p}(x+2\eta e^{\sqrt{-1}\theta}))}} \frac{d\theta}{2\pi}.$$

Hence, by Markov's inequality,

$$\mathbf{P}\Big(\max_{z\in\partial B(x,\eta)}\big|\bar{p}(z)\big|\geq t\Big)$$
$$\leq \mathbf{E}\bigg(\exp\bigg(\int_0^{2\pi}\frac{|\bar{p}(x+2\eta e^{\sqrt{-1}\theta})|}{10\sqrt{\mathbf{Var}(\bar{p}(x+2\eta e^{\sqrt{-1}\theta}))}}\frac{d\theta}{2\pi}\bigg)^2\bigg)e^{-t^2/(10^2CL^{2\rho+1-4c_2})}.$$

Applying Jensen's inequality for convex functions $x\to x^2$ and $x\to e^x$ and Fubini's theorem gives

$$\mathbf{E}\bigg(\exp\bigg(\int_0^{2\pi}\frac{|\bar{p}(x+2\eta e^{\sqrt{-1}\theta})|}{10\sqrt{\mathbf{Var}(\bar{p}(x+2\eta e^{\sqrt{-1}\theta}))}}\frac{d\theta}{2\pi}\bigg)^2\bigg)$$
$$\leq \int_0^{2\pi}\mathbf{E}\exp\bigg(\frac{|\bar{p}(x+2\eta e^{\sqrt{-1}\theta})|^2}{100\mathbf{Var}(\bar{p}(x+2\eta e^{\sqrt{-1}\theta}))}\bigg)\frac{d\theta}{2\pi}.$$

Let $z=x+2\eta e^{\sqrt{-1}\theta}$ then the real part and imaginary part of $\frac{\bar{p}(z)}{\sqrt{\mathbf{Var}(\bar{p}(z))}}=:$ $X_z+\sqrt{-1}Y_z$ are normally distributed with mean 0 and variance at most 1. Hence, by Cauchy–Schwarz inequality,

$$\mathbf{E}e^{10^{-2}|X_z+\sqrt{-1}Y_z|^2}=\mathbf{E}e^{10^{-2}X_z^2}e^{10^{-2}Y_z^2}\leq \mathbf{E}e^{2.10^{-2}X_z^2}+\mathbf{E}e^{2.10^{-2}Y_z^2}\leq C.$$

This completes the proof of (5.5). Now set

$$t=L^{\rho+1/2-2c_2+c_3}, \tag{5.6}$$

then (5.5) becomes

$$\mathbf{P}\bigg(\max_{z\in\partial B(x,\eta)}\big|p(z)-\mathbf{E}p(z)\big|\geq\frac{1}{2}t\bigg)\leq Ce^{-t^2/(4CL^{2\rho+1-4c_2})}\leq\delta^{2c_3}\gamma^{3/2}. \tag{5.7}$$

(To prove Lemma 5.1 for Q, we set $t=L^{1/2-2c_2+c_3}$.)

Let $A_2=\{\omega:\max_{z\in\partial B(x,\eta)}|p(z)-\mathbf{E}p(z)|\geq\frac{1}{2}t\}$.

Now, since g is a linear with real coefficients [$P(x)$ and $P'(x)$], one has

$$\min_{z\in\partial B(x,\eta)}\big|g(z)\big|=\min\big|g(x\pm\eta)\big|.$$

And so,

$$\mathbf{P}\Big(\min_{z\in\partial B(x,\eta)}\big|g(z)\big|\leq t\Big)\leq\mathbf{P}\big(\big|g(x+\eta)\big|\leq t\big)+\mathbf{P}\big(\big|g(x-\eta)\big|\leq t\big).$$

Since $g(x\pm\eta)$ is normally distributed,

$$\begin{aligned}\mathbf{P}\big(\big|g(x\pm\eta)\big|\leq t\big)&\leq\mathbf{P}\big(\big|g(x\pm\eta)-\mathbf{E}g(x\pm\eta)\big|\leq t\big)\\&\leq\frac{t}{\sqrt{\mathbf{Var}(g(x\pm\eta))}}=\frac{t}{|v_x\pm\eta v_x'|}.\end{aligned} \tag{5.8}$$

Using Condition 1, we have

$$\eta|v_x'| \le C\eta\sqrt{\sum_{i=0}^{n} i^{2\rho+2}x^{2i}} \le C\eta L^{\rho+3/2} = CL^{\rho+1/2-c_2}, \tag{5.9}$$

$$|v_x| \ge \frac{1}{C}\sqrt{\sum_{i=L/40}^{L/20} i^{2\rho}x^{2i}} \ge \frac{1}{C}L^{\rho+1/2}, \tag{5.10}$$

which together give $|v_x - \eta v_x'| \ge \frac{1}{C}L^{\rho+1/2}$.

[To prove Lemma 5.1 for Q observe that $|\frac{d_i}{d_0}| \ge \frac{1}{C}$ for all $i \le \frac{n}{2}$, and hence for all $i \le L/20$; therefore, by the same argument as above, for the vector field $v_z = (\frac{d_i}{d_0}z^i)_{i=0}^n$, one has $|v_x| \ge \frac{1}{C}L^{1/2}$, which is again similar to the case $\rho = 0$ for P; now for $\eta|v_x'|$ we similarly have

$$\begin{aligned}
\eta|v_x'| &\le C\eta\sqrt{\sum_{i=0}^{n} i^2\big[(n-i)/n\big]^{2\rho}x^{2i}} \\
&\le C\eta\sqrt{\sum_{i\le n/2} i^2x^{2i} + n^{2-2\rho}x^{n/2}\sum_{i>n/2}(n-i)^{2\rho}} \\
&\le C\eta\sqrt{L^3 + n^3x^{n/2}} \le C\eta L^{3/2} = CL^{1/2-c_2}
\end{aligned}$$

which is similar to the case $\rho = 0$ for P.]

And so, by (5.6), the bound (5.8) becomes

$$\begin{aligned}
\mathbf{P}\big(\big|g(x\pm\eta)\big| \le t\big) &\le \mathbf{P}\big(\big|g(x\pm\eta) - \mathbf{E}g(x\pm\eta)\big| \le t\big) \\
&\le CL^{-2c_2+c_3} \le CL^{-3/2c_2-2c_3} = C\delta^{2c_3}\gamma^{3/2}.
\end{aligned} \tag{5.11}$$

Hence,

$$\mathbf{P}\Big(\min_{z\in\partial B(x,\eta)}\big|g(z)\big| \le t\Big) \le C\delta^{2c_3}\gamma^{3/2}. \tag{5.12}$$

Let $A_3 = \{\omega : \min_{z\in\partial B(x,\eta)}|g(z)| \le t\}$, and $A_4 = A_1 \setminus (A_2 \cup A_3)$

If Condition 2(2a) holds, that is, $\mathbf{E}\xi_i = 0$ for all $N_0 \le i \le n$, then by (5.3),

$$\big|\mathbf{E}p(z)\big| \le \eta^2\sum_{i=0}^{N_0}|\mathbf{E}\xi_i||c_i|i^2(1+3/n)^n \le C\eta^2 \le \frac{t}{2}$$

for every $z \in \partial B(x,\eta)$. This together with (5.7) give

$$\mathbf{P}\Big(\max_{z\in\partial B(x,\eta)}\big|p(z)\big| \ge t\Big) \le \delta^{2c_3}\gamma^{3/2}.$$

And so $\mathbf{P}(A_1) \le \mathbf{P}(A_3) + \mathbf{P}(\max_{z\in\partial B(x,\eta)}|p(z)| \ge t) \le \delta^{2c_3}\gamma^{3/2}$ as desired.

[Similarly, for Q, one has $|\mathbf{E}p(z)| \le \eta^2 \sum_{i=0}^{N_0} |\mathbf{E}\xi_i| \frac{|c_i|}{|c_n|}(n-i)^2(1-\frac{1}{2L})^{n-i} \le C\eta^2 n^{2-\rho} e^{-n/2L} \le C\eta^2 L^{2-\rho} = CL^{-\rho-2c_2} \le \frac{t}{2}$ for every $z \in \partial B(x,\eta)$ because $\rho > -1/2$.]

Now, if Condition 2(2b) holds, and $x < 0$, that is, x is in $-I(\delta) + (-10^{-4}\delta, 10^{-4}\delta)$. Recall that $\rho \ge 0$ under Condition 2(2b). Then for every $z \in \partial B(x,\eta)$,

$$\begin{aligned}\left|\mathbf{E}p(z)\right| &\le C\eta^2 + |\mu| \left| \sum_{i=N_0}^{n} c_i z^i - \sum_{i=N_0}^{n} c_i x^i - (z-x) \sum_{i=N_0}^{n} i c_i x^{i-1} \right| \\ &\le C\eta^2 + C\eta^2 \max_{z' \in \partial B(x,\eta)} \left| \sum_{i=0}^{n} \mathfrak{P}(i) i(i-1) z'^{i-2} \right|,\end{aligned}$$

in which we used the fact that the contributions of the sums from $i = 0$ to $i = N_0 - 1$ are just $O(\eta^2)$ as showed in the case of Condition 2(2a). Observe that $\mathfrak{P}(i)i(i-1) = \sum_{j=0}^{\rho+2} e_j i(i-1)\cdots(i-j+1)$ for some constants e_j, and for each $0 \le j \le \rho+2$,

$$\begin{aligned}\left| \sum_{i=0}^{n} i(i-1)\cdots(i-j+1) z'^{i-j} \right| &= \left| \left(\frac{1-z'^{n+1}}{1-z'} \right)^{(j)} \right| \\ &\le C n^j |z'|^{n-j+1} \le C L^j \frac{n^j}{L^j} e^{-n/L} \\ &\le C L^j \le C L^{\rho+2},\end{aligned}$$

where in the first inequality, we used the bounds $|1-z'| \ge |1-x| - |x-z'| \ge 1$.

This shows that $|\mathbf{E}p(z)| \le CL^{\rho-2c_2}|\mu| \le \frac{t}{2}$. From this, the same proof as for Condition 2(2a) applies.

[Similarly, for Q, one has

$$\begin{aligned}\left|\mathbf{E}p(z)\right| &\le C\eta^2 L^2 + C\eta^2 \max_{z' \in \partial B(x,\eta)} \left| \sum_{i=0}^{n} \frac{\mathfrak{P}(n-i)}{\mathfrak{P}(n)} i(i-1) z'^{i-2} \right| \\ &= O(\eta^2 L^2) = O(L^{-2c_2}) \le \frac{t}{2}\end{aligned}$$

in which, again, we used the fact that the contribution of the sums from $i = n - N_0$ to $i = n$ is $O(\eta^2 L^2)$ as showed in the case of Condition 2(2a).]

Now, if Condition 2(2b) holds, and $x \ge 0$, that is, x is in $I(\delta) + (-10^{-4}\delta, 10^{-4}\delta)$. Without loss of generality, assume that $\mu \ge 0$ and $c_i > 0$ for all i sufficiently large, say $i \ge N_0$ (by replacing c_i by $-c_i$ and ξ_i by $-\xi_i$ if needed).

We have

$$\mathbf{P}(A_1) \le \mathbf{P}(A_2) + \mathbf{P}(A_3) + \mathbf{P}(A_4) \le C\delta^{2c_3}\gamma^{3/2} + \mathbf{P}(A_4). \tag{5.13}$$

If $|\mathbf{E}p(z)| \le \frac{t}{2}$ for every $z \in \partial B(x,\eta)$, then as in the above case we also have $\mathbf{P}(A_1) \le \delta^{2c_3}\gamma^{3/2}$. Otherwise, assume that there exists $z_0 \in B(x,\delta)$ such that

$|\mathbf{E}p(z_0)| > \frac{t}{2}$. Without loss of generality, we choose z_0 that maximizes $|\mathbf{E}p(z_0)|$ in that (closed) ball. Let $m(z) = \mathbf{E}P(z) = \sum_{i=0}^{n} c_i \mathbf{E}\xi_i z^i$. Then

$$\begin{aligned}|\mathbf{E}p(z_0)| = |m(z_0) - m(x) - m'(x)(z_0 - x)| &\leq \frac{|z_0 - x|^2}{2} \max_{z\in\partial B(x,\eta)} |m''(z)| \\ &\leq o(t) + \mu\eta^2 \sum_{i=0}^{i=n} c_i i(i-1)(x+\eta)^{i-2}.\end{aligned}$$

[For P the $o(t)$ is $C\eta^2$, and for Q the $o(t)$ is $C\eta^2 L^2$.]

Observe by a similar bound as in (5.3) that

$$\left|\sum_{i=n\wedge 2(4+\rho)L\log L}^{n} c_i i(i-1)(x+\theta z_0)^{i-2}\right| \leq C L^{\rho+3} \int_{2(4+\rho)\log L}^{\infty} e^{-x/2}\,dx \leq 1.$$

Hence,

$$\begin{aligned}\frac{t}{2} < |\mathbf{E}p(z_0)| &\leq o(t) + \eta^2 \sum_{i=0}^{i=n\wedge 2(4+\rho)L\log L} c_i i^2 (x+\eta)^{i-2} \\ &\leq 2\eta^2 \sum_{i=0}^{i=n\wedge 2(4+\rho)L\log L} c_i i^2 (x+\eta)^{i-2}.\end{aligned}$$

Now,

$$\begin{aligned}m(x) &\geq \mu \sum_{i=0}^{i=n\wedge 2(4+\rho)L\log L} c_i x^i - o(t) \\ &\geq \frac{1}{C}\mu L^{-2}\log^{-2} L \sum_{i=0}^{i=n\wedge 2(4+\rho)L\log L} c_i i^2 (x+\eta)^i - o(t) \\ &\geq \frac{1}{C}\frac{L^{2c_2}}{\log^2 L}\eta^2 \sum_{i=0}^{i=n\wedge 2(4+\rho)L\log L} c_i i^2 (x+\eta)^{i-2} - o(t) \\ &\geq \frac{1}{C}\frac{L^{2c_2}}{\log^2 L}|\mathbf{E}p(z_0)| \geq \frac{1}{C}\frac{L^{2c_2}}{\log^2 L} t = \frac{1}{C}\frac{L^{\rho+1/2+c_3}}{\log^2 L}.\end{aligned} \tag{5.14}$$

Similarly,

$$\begin{aligned}\eta m'(x) &\leq C + \eta \sum_{i=0}^{i=n\wedge 2(4+\rho)L\log L} c_i i x^{i-1} \\ &\leq C + C\eta L(\log L) m(x) \leq C\frac{\log L}{L^{c_2}} m(x) \leq \frac{m(x)}{2}.\end{aligned}$$

Thus,

$$\mathbf{E}g(x \pm \eta) = m(x) \pm \eta m'(x) \geq \frac{m(x)}{2}. \tag{5.15}$$

By this and (5.9) and its analog for v_x show that

$$\sqrt{\mathbf{Var}\, g(x \pm \eta)} \leq C L^{\rho+1/2} \leq \frac{\mathbf{E}g(x \pm \eta)}{L^{c_3/2}}. \tag{5.16}$$

On the event A_4, we know that $\min|g(x \pm \eta)| \leq \max_{z \in \partial B(x,\eta)} |p(z)|$. Choose any z in the closed ball $cl(B(x,\eta))$ that maximizes $|p|$. Then $\min|g(x \pm \eta)| \leq |p(z)|$. Since $A_4 \cap A_2 = \varnothing$, $|p(z)| \leq |\mathbf{E}p(z_0)| + t/2 \leq 2|\mathbf{E}p(z_0)|$. Then, by (5.14) and (5.15),

$$\min\left|g(x \pm \eta)\right| \leq \left|p(z)\right| \leq C\frac{\log^2 L}{L^{2c_2}} \min\left|\mathbf{E}g(x \pm \eta)\right| \leq \frac{1}{2} \min \mathbf{E}g(x \pm \eta). \tag{5.17}$$

Finally, by (5.16), we have

$$\begin{aligned}
&\mathbf{P}\left(\left|g(x \pm \eta)\right| \leq \frac{1}{2}\mathbf{E}g(x \pm \eta)\right) \\
&\quad \leq \mathbf{P}\left(\frac{|g(x \pm \eta) - \mathbf{E}g(x \pm \eta)|}{\sqrt{\mathbf{Var}(g(x \pm \eta))}} \geq \frac{\mathbf{E}g(x \pm \eta)}{2\sqrt{\mathbf{Var}(g(x \pm \eta))}}\right) \\
&\quad \leq \mathbf{P}\left(\left|N(0,1)\right| \geq \frac{L^{c_3/2}}{2}\right) \leq \delta^{2c_3}\gamma^{3/2}.
\end{aligned}$$

This proves (5.17), and thus, $\mathbf{P}(A_4) \leq C\delta^{2c_3}\gamma^{3/2}$. So is A_1. □

Now, for every $1 \leq i \leq k$, consider the strip $S_i = [\check{x}_i - r_0, \check{x}_i + r_0] \times [-\gamma/4, \gamma/4]$. We can cover S by $O(\gamma^{-1})$ balls of the form $B(\check{x}, \gamma)$ where $x \in [\check{x}_i - r_0, \check{x}_i + r_0]$. Using Lemma 5.1, we obtain

$$\begin{aligned}
&\mathbf{P}(\text{there is at least 1 (or equivalently 2) root in } S_i \setminus \mathbb{R}) \\
&\quad = O\left(\gamma^{-1}\gamma^{3/2}\right) = O\left(\gamma^{1/2}\right).
\end{aligned} \tag{5.18}$$

Consider $\hat{F}_i(z) = F_i(\mathrm{Re}(z))\phi(\frac{4\,\mathrm{Im}(z)}{\gamma})$, where ϕ is a bump function on $\mathbb{R}$ that is supported on $[-1, 1]$ and is 1 at 0. Then $\hat{F}_i$ is a smooth function supported on $S_i - \check{x}_i$ and $|\hat{F}_i| \leq 1$, and $|\nabla^a \hat{F}_i| = O(\gamma^{-a})$ for $0 \leq a \leq 3$.

Set $X_{\check{x}_i, \hat{F}_i} = \sum_{j=1}^n \hat{F}_i(\zeta_j^{\check{P}} - \check{x}_i)$ and

$$D_{\check{x}_i, F_i} = X_{\check{x}_i, \hat{F}_i} - X_{x_i, F_i, \mathbb{R}} = \sum_{\zeta_i^{\check{P}} \notin \mathbb{R}} \hat{F}_i\left(\zeta_i^{\check{P}} - \check{x}_i\right).$$

Observe that $|D_{\check{x}_i, F_i}| \leq N_{\check{P}} B(\check{x}_i, 2r_0)$, and from (5.18), $D_{\check{x}_i, F_i} = 0$ with probability at least $1 - O(\gamma^{1/2})$.

Let ϕ_0 be a bump function supported on $B(0, 4r_0)$ that equals 1 on $B(0, 2r_0)$ and $|\triangle^a \phi_0| \le C$ for all $0 \le a \le 3$, then

$$\max\{|X_{\check{x}_i,\hat{F}_i}|, |X_{x_i,F_i,\mathbb{R}}|, |D_{\check{x}_i,F_i}|\} \le \sum_{j=1}^n \phi_0(\zeta_j^{\check{P}} - \check{x}_i) =: X_{\check{x}_i,\phi_0}.$$

Let $c_4 = \frac{c_2}{4(k+l)^2}$. By Proposition 4.6, $N_{\check{P}} B(\check{x}_i, 2r_0) = N_P B(x_i, 2r_0 10^{-3}\delta) \le \delta^{-c_4}$ with probability at least $1 - C\gamma(\delta)$. And from Section 4.4, we have $\mathbf{E}(|X_{\check{x}_i,\phi_0}|^{k+l} \mathbf{1}_{N_{\check{P}} B(\check{x}_i,2r_0) > \delta^{-c_4}}) \le C\delta^{1/22}$.

Hence,

$$\begin{aligned}
&\mathbf{E}|X_{\check{x}_i,\hat{F}_i} - X_{\check{x}_i,F_i,\mathbb{R}}|^{k+l} \\
&\quad = \mathbf{E}(|D_{\check{x}_i,F_i}|^{k+l} \mathbf{1}_{N_{\check{P}} B(\check{x}_i,2r_0) \le \delta^{-c_4}}) + \mathbf{E}(|D_{\check{x}_i,F_i}|^{k+l} \mathbf{1}_{N_{\check{P}} B(\check{x}_0,2r_0) > \delta^{-c_4}}) \\
&\quad \le C\delta^{-c_4(k+l)} \gamma^{1/2} + \mathbf{E}(|X_{\check{x}_i,\phi_0}|^{k+l} \mathbf{1}_{N_{\check{P}} B(\check{x}_i,2r_0) > \delta^{-c_4}}) \le C\delta^{c_4(k+l)^2}.
\end{aligned}$$

Moreover, by another application of Proposition 4.6 and Section 4.4, we obtain

$$\begin{aligned}
&\max\{\mathbf{E}|X_{\check{x}_i,\hat{F}_i}|^{k+l}, \mathbf{E}|X_{\check{x}_i,F_i,\mathbb{R}}|^{k+l}\} \\
&\quad = \mathbf{E}(|X_{\check{x}_i,\phi_0}|^{k+l} \mathbf{1}_{N_{\check{P}} B(\check{x}_i,4r_0) \le \delta^{-c_4}}) + \mathbf{E}(|X_{\check{x}_i,\phi_0}|^{k+l} \mathbf{1}_{N_{\check{P}} B(\check{x}_0,4r_0) > \delta^{-c_4}}) \\
&\quad \le C\delta^{-c_4(k+l)} + C\delta^{1/22} \le C\delta^{-c_4(k+l)}.
\end{aligned}$$

Similarly, for each $1 \le j \le l$, let $\hat{G}_j(z) = G_j(z)\eta(\mathrm{Im}(z + \check{z}_j)/\gamma)$ where η is a bump function on $\mathbb{R}$ supported on $[1/2, \infty)$ and equal 1 on $[1, \infty)$. And let $X_{\check{z}_j,\hat{G}_j} = \sum_{i=1}^n \hat{G}_j(\zeta_i^{\check{P}} - \check{z}_j)$. Then $\mathbf{E}|X_{\check{z}_j,\hat{G}_j} - X_{\check{z}_j,G_j,\mathbb{C}_+}|^{k+l} \le C\delta^{c_4(k+l)^2}$ and $\max\{\mathbf{E}|X_{\check{z}_j,\hat{G}_j}|^{k+l}, \mathbf{E}|X_{\check{z}_j,G_j,\mathbb{C}_+}|^{k+l}\} \le C\delta^{-c_4(k+l)}$.

By telescoping the difference and applying Hölder's inequality, we obtain

$$\mathbf{E}\left|\left(\prod_{i=1}^k X_{\check{x}_i,F_i,\mathbb{R}}\right)\left(\prod_{j=1}^l X_{\check{z}_j,G_j,\mathbb{C}_+}\right) - \left(\prod_{i=1}^k X_{\check{x}_i,\hat{F}_i}\right)\left(\prod_{j=1}^l X_{\check{z}_j,\hat{G}_j}\right)\right| \le C\delta^{c_4}.$$

Combining this with (5.2) with H_j's being $\hat{F}_i/O(\gamma^{-3})$ and $\hat{G}_j/O(\gamma^{-3})$, respectively, we get the desired result.

6. Proof of Lemma 2.5 and Corollary 2.6.

PROOF OF LEMMA 2.5. If suffices to show that

$$\mathbf{E} N_{P_n}\left(\left[-1 + \frac{1}{C}, 1 - \frac{1}{C}\right]\right) \le M(C), \tag{6.1}$$

and $\mathbf{E} N_{Q_n}([-1 + \frac{1}{C}, 1 - \frac{1}{C}]) \le M(C)$, for some constant $M(C)$.

Again, the proof for the second inequality is the same as the first. We follow the approach in [15]. By (4.27), there exist constants d and q such that $\mathbf{P}(|\xi_i| \le d) \le q < 1$ for all i.

For $k \ge N_0$, let $B_k = \{\omega : |\xi_{N_0}| \le d, \dots, |\xi_{k-1}| \le d, |\xi_k| > d\}$. Then $\mathbf{P}(B_k) \le q^{k-N_0}$.

By mean value theorem and Jensen's inequality, we have

$$N_P\left[-1+\frac{1}{C}, 1-\frac{1}{C}\right] \le k + N_{P^{(k)}}\left[-1+\frac{1}{C}, 1-\frac{1}{C}\right] \le k + \frac{\log \frac{M}{|P^{(k)}(0)|}}{\log \frac{R}{r}},$$

where $R = 1 - \frac{1}{2C}$, $r = 1 - \frac{1}{C}$, and $M = \sup_{|z|=R} |P^{(k)}(z)|$. On B_k, we have

$$N_P\left[-1+\frac{1}{C}, 1-\frac{1}{C}\right] \le k + \frac{\log \frac{\sum_{j=k}^n c_{jk} \frac{|c_j|}{|c_k|} |\xi_j|}{d}}{\log \frac{R}{r}},$$

where $c_{jk} = j(j-1)\cdots(j-k+1)R^{j-k}/k!$. And so,

$$\begin{aligned}\mathbf{E}N_P\left[-1+\frac{1}{C}, 1-\frac{1}{C}\right] &\le \sum_{k=N_0}^{n+1} k\mathbf{P}(B_k) + \frac{\log 1/d}{\log \frac{R}{r}} \sum_{k=N_0}^{n+1} \mathbf{P}(B_k) \\ &\quad + \frac{1}{\log \frac{R}{r}} \sum_{k=N_0}^{n+1} \int_{B_k} \log\left(\sum_{j=k}^n c_{jk} \frac{|c_j|}{|c_k|} |\xi_j|\right) d\mathbf{P}.\end{aligned}$$

Thus, to show (6.1), it suffices to show that

$$\sum_{k=N_0}^{n+1} \int_{B_k} \log(R_k)\, d\mathbf{P} \le C', \tag{6.2}$$

for some $C' = C'(C)$, where $R_k = (\rho+1+k)^{-\rho-1} \sum_{j=k}^n c_{jk} \frac{|c_j|}{|c_k|} |\xi_j|$. Then

$$\mathbf{E}R_k \le C'(\rho+1+k)^{-\rho-1} \sum_{j=k}^n \binom{j}{k} \frac{|c_j|}{|c_k|} R^{j-k} \le \frac{C' k^{-\rho}}{(1-R)^{k+\rho+1}}.$$

Let $B_{ki} = \{\omega \in B_k : e^i \mathbf{E}R_k < R_k \le e^{i+1}\mathbf{E}R_k\}$. Then $\mathbf{P}(B_{ki}) \le e^{-i}$ by Markov's inequality. Let $i_0 = \lfloor -\log q^k \rfloor$, then

$$\begin{aligned}\int_{B_k} \log R_k\, d\mathbf{P} &\le \mathbf{P}(B_k)\log(e^{i_0}\mathbf{E}R_k) + \sum_{i=i_0}^{\infty} \int_{B_{ki}} \log R_k\, d\mathbf{P} \\ &\le C' q^k (k+2+\rho - k\log q - \rho \log k).\end{aligned}$$

This proves (6.2) and completes the proof. □

PROOF OF COROLLARY 2.6. Let C be the constant in Theorem 2.4 with $k=1$. As a consequence of the above lemma, we only need to concentrate on the domain $\mathbb{R}\cap A(0,1-\frac{1}{C},1+\frac{1}{C})$.

Let c be the constant in Theorem 2.4, and let $\alpha=c/7$. We will prove that for every $\frac{1}{20n}\le\delta\le\frac{1}{C}$ and real number $x_0\in\mathbb{R}$ such that $|x_0|\in I(\delta)$, we have

$$\begin{aligned}&\big|\mathbf{E}N_P\big(x_0-10^{-7}\delta,x_0+10^{-7}\delta\big)-\mathbf{E}N_{\widetilde{P}}\big(x_0-10^{-7}\delta,x_0+10^{-7}\delta\big)\big|\\&\qquad=O\big(\delta^{\alpha/2}\big),\end{aligned}\tag{6.3}$$

and when $\frac{1}{10n}\le\delta\le\frac{1}{C}$,

$$\begin{aligned}&\big|\mathbf{E}N_P\big(x_0-10^{-7}\delta,x_0+10^{-7}\delta\big)-\mathbf{E}N_{\widetilde{P}}\big(x_0-10^{-7}\delta,x_0+10^{-7}\delta\big)\big|\\&\qquad=O\big(\delta^{\alpha/2}\big).\end{aligned}\tag{6.4}$$

From (6.3), we can conclude that $|\mathbf{E}N_P(\pm I(\delta))-\mathbf{E}N_{\widetilde{P}}(\pm I(\delta))|=O(\delta^{\alpha/2})$ for all $\frac{1}{20n}\le\delta\le\frac{1}{C}$. Letting $\delta=\frac{1}{20n},\frac{1}{10n},\dots,\frac{2^m}{20n}$ where $\frac{2^{m-1}}{20n}<\frac{1}{C}\le\frac{2^m}{20n}$ and applying triangle inequality, we obtain

$$\left|\mathbf{E}N_P\left(\pm\left(1-\frac{2^{m+1}}{n},1+\frac{1}{n}\right)\right)-\mathbf{E}N_{\widetilde{P}}\left(\pm\left(1-\frac{2^{m+1}}{n},1+\frac{1}{n}\right)\right)\right|=O(1).$$

This together with the analogue for Q give the desired result. (By definition of Q, we have $\mathbf{E}N_Q[a,b]=\mathbf{E}N_P[1/b,1/a]$ if $0\le a<b\le\infty$ or $-\infty\le a<b\le 0$.)

As for the proof of (6.3), let $\frac{1}{20n}\le\delta\le\frac{1}{C}$ and let x_0 be a real number with $|x|\in I(\delta)$.

Let G be a smooth function supported on $[-10^{-4}-\delta^\alpha,10^{-4}+\delta^\alpha]$ such that $0\le G\le 1$, $G=1$ on $[-10^{-4},10^{-4}]$, and $\|\nabla^a G\|\le C\delta^{-6\alpha}$ for all $0\le a\le 6$. We have

$$\begin{aligned}&\mathbf{E}N_P\big[x_0-10^{-7}\delta,x_0+10^{-7}\delta\big]\\&\quad=\mathbf{E}N_{\check{P}}\big[\check{x}_0-10^{-4},\check{x}_0+10^{-4}\big]\le\mathbf{E}\sum_{\zeta_i^{\check{P}}\in\mathbb{R}}G\big(\zeta_i^{\check{P}}-\check{x}_0\big)\\&\quad\le\mathbf{E}\sum_{\zeta_i^{\check{\tilde{P}}}\in\mathbb{R}}G\big(\zeta_i^{\check{\tilde{P}}}-\check{x}_0\big)+C\delta^{c-6\alpha}\qquad\text{by Theorem 2.4}\\&\quad\le\mathbf{E}\sum_{i=1}^n\mathbf{1}_{[-\delta^\alpha-10^{-4},\delta^\alpha+10^{-4}]}\big(\zeta_i^{\check{\tilde{P}}}-\check{x}_0\big)+C\delta^{c-6\alpha}\\&\quad\le\mathbf{E}N_{\widetilde{P}}\big[x_0-10^{-7}\delta,x_0+10^{-7}\delta\big]+\mathcal{I}_{\widetilde{P}}+C\delta^\alpha,\end{aligned}$$

where $\mathcal{I}_{\widetilde{P}}=\mathbf{E}\sum_{i=1}^n\mathbf{1}_{\pm[10^{-7}\delta,10^{-7}\delta+10^{-3}\delta^{\alpha+1}]}(\zeta_i^{\widetilde{P}}-x_0)$. We will show later that $\mathcal{I}_{\widetilde{P}}=O(\delta^{\alpha/2})$.

Thus,

$$\mathbf{E}N_P[x_0 - 10^{-7}\delta, x_0 + 10^{-7}\delta] \le \mathbf{E}N_{\widetilde{P}}[x_0 - 10^{-7}\delta, x_0 + 10^{-7}\delta] + C\delta^{\alpha/2}.$$

By similar arguments with the function G being replaced by one supported on $[-10^{-4}, 10^{-4}]$ such that $0 \le G \le 1$ and $G = 1$ on $[-10^{-4} + \delta^\alpha, 10^{-4} - \delta^\alpha]$, we have

$$\mathbf{E}N_P[x_0 - 10^{-7}\delta, x_0 + 10^{-7}\delta] \ge \mathbf{E}N_{\widetilde{P}}[x_0 - 10^{-7}\delta, x_0 + 10^{-7}\delta] - C\delta^{\alpha/2}.$$

This gives (6.3) for P. Hence, to finish, we only need to prove the stated bound on $\mathcal{I}_{\widetilde{P}}$ and $\mathcal{I}_{\widetilde{Q}}$. Let $[a, b] = x_0 \pm [10^{-7}\delta, 10^{-7}\delta + 10^{-3}\delta^{\alpha+1}]$. By a Kac–Rice-type formula (see, for instance, [12], Theorem 2.5), one has

$$\mathbf{E}N_{\widetilde{P}}[a, b] \le \int_a^b \sqrt{\frac{\mathcal{S}}{\mathcal{P}^2}}\, dt + \int_a^b \frac{|m'|\mathcal{P} + |m|\mathcal{R}}{\mathcal{P}^{3/2}} e^{-\frac{1}{2}(\frac{m}{\sqrt{\mathcal{P}}})^2}\, dt, \tag{6.5}$$

for any $a \le b$, where $m(t) = \mathbf{E}\widetilde{P}(t)$, $\mathcal{P} = \mathbf{Var}(\widetilde{P}) = \sum_{i=0}^n c_i^2 t^{2i}$, $\mathcal{Q} = \mathbf{Var}(\widetilde{P}') = \sum_{i=0}^n c_i^2 i^2 t^{2i-2}$, $\mathcal{R} = \mathbf{Cov}(\widetilde{P}, \widetilde{P}') = \sum_{i=0}^n c_i^2 i t^{2i-1}$, and $\mathcal{S} = \mathcal{P}\mathcal{Q} - \mathcal{R}^2 = \sum_{i<j} (j-i)^2 c_i^2 c_j^2 t^{2i+2j-2}$.

First, we will bound the second integral. By similar bounds as in (5.4) and (5.10), we have for every $t \in [a, b]$,

$$\mathcal{P} \ge \frac{1}{C}\delta^{-2\rho-1}, \qquad \mathcal{R} \le C\delta^{-2\rho-2} \le C\delta^{-1}\mathcal{P}.$$

Additionally, by the same argument as in the proof of Theorem 2.4 [more precisely Lemma 5.1 near the estimate (5.14)], one can show that under Condition 2(2), for every $t \in [a, b]$, $|m'(t)| \le C\delta^{-\rho-1} + C\delta^{-1}\log\frac{1}{\delta}|m(t)|$. Thus,

$$\frac{|m'|\mathcal{P} + |m|\mathcal{R}}{\mathcal{P}^{3/2}} e^{-\frac{1}{2}(\frac{m}{\sqrt{\mathcal{P}}})^2} \le C\delta^{-1/2} + C\delta^{-1}\log\frac{1}{\delta}\frac{|m|}{\sqrt{\mathcal{P}}} e^{-\frac{1}{2}(\frac{m}{\sqrt{\mathcal{P}}})^2} \le C\delta^{-1}\log\frac{1}{\delta},$$

where in the last inequality, we used the boundedness of the function $x \to xe^{-x^2/2}$ on $\mathbb{R}$. Since the length of the interval $[a, b]$ is $C\delta^{\alpha+1}$, the second integral in (6.5) is of order $O(\delta^{\alpha/2})$ as desired.

Hence, it remains to bound the first integral in (6.5). By symmetry, we may assume that $a > 0$. We first reduce to the hyperbolic polynomials for which that integral is easier to handle. Consider the corresponding hyperbolic polynomials with coefficients $c_i^{\text{hyper}} = \sqrt{\frac{(2\rho+1)\cdots(2\rho+i)}{i!}}$. A routine estimation shows that $\frac{1}{C}i^\rho \le c_i^{\text{hyper}} \le Ci^\rho$ for some constant C. And thus, by condition (2.5), $\frac{1}{C'}c_i^{\text{hyper}} \le |c_i| \le C'c_i^{\text{hyper}}$ for all $N_0 \le i$, and so when $|t| \ge \frac{1}{2}$, one has $\mathcal{S}(t) \le C'\mathcal{S}^{\text{hyper}}(t)$ and $\mathcal{P}(t) \ge \frac{1}{C'}\mathcal{P}^{\text{hyper}}(t)$. Thus, $\sqrt{\frac{\mathcal{S}}{\mathcal{P}^2}} \le C'\sqrt{\frac{\mathcal{S}^{\text{hyper}}}{(\mathcal{P}^{\text{hyper}})^2}}$.

If $\frac{1}{2} \le t \le 1 - \frac{(100\rho+100)\log n}{n}$, one has

$$\mathcal{P}^{\text{hyper}}(t) = \frac{1}{(1-t^2)^{2\rho+1}} - \sum_{i=n+1}^{\infty} \frac{(2\rho+1)\cdots(2\rho+i)}{i!} t^{2i},$$

and the last term is bounded from above by $\sum_{i=1}^{\infty} \frac{(2\rho+1)\cdots(2\rho+i)}{i!} t^{2i} A_i$ where

$$\begin{aligned} A_i &= \frac{(2\rho+i+1)\cdots(2\rho+i+n)}{(i+1)\cdots(i+n)} t^{2n} \\ &\le \frac{(i+n+1)\cdots(i+n+\lceil 2\rho\rceil)}{(i+1)\cdots(i+\lceil 2\rho\rceil)} t^{2n} = o(n^{-100\rho-100}). \end{aligned}$$

Thus, $\mathcal{P}^{\text{hyper}} = \frac{1}{(1-t^2)^{2\rho+1}}(1+o(n^{-100\rho-100}))$. Similarly,

$$\begin{aligned} \mathcal{Q} &= \left(\frac{(2\rho+1)(2\rho+2)t^2}{(1-t^2)^{(2\rho+3)}} + \frac{2\rho+1}{(1-t^2)^{2\rho+2}}\right)[1+o(n^{-100\rho-100})], \\ \mathcal{R} &= \frac{(2\rho+1)t}{(1-t^2)^{2\rho+2}}[1+o(n^{-100\rho-100})], \end{aligned}$$

therefore,

$$\sqrt{\frac{\mathcal{S}^{\text{hyper}}}{(\mathcal{P}^{\text{hyper}})^2}} = \frac{\sqrt{2\rho+1}}{1-t^2}(1+O(n^{-12\rho-12})).$$

Plugging into (6.5) with $[a,b] = x_0 \pm [10^{-7}\delta, 10^{-7}\delta + 10^{-3}\delta^{\alpha+1}]$ gives the desired bound for $\delta \ge (200\rho+200)n^{-1}\log n$.

Next, if $1 + \frac{2}{n} \ge t \ge 1 - \frac{(500\rho+500)\log n}{n}$, we will prove that

$$\frac{\mathcal{S}}{\mathcal{P}^2} \le \frac{O(n)}{|1-t|}. \tag{6.6}$$

This together with (6.5) will give the desired bound for $\delta \le \frac{(200\rho+200)\log n}{n}$.

To prove (6.6), observe that $\mathcal{S} \le 4\sum_{0\le i<j\le n}(j-i)^2 c_i^2 c_j^2 t^{2i+2j}$. Set $M = \frac{1}{|1-t|}$. We have

$$\sum_{0\le i<j\le n\wedge(i+\sqrt{nM})} (j-i)^2 c_i^2 c_j^2 t^{2i+2j} \le \frac{n}{|1-t|}\mathcal{P}^2(t). \tag{6.7}$$

And so, we only need to work on the summands corresponding to $0 \le i \le i + \sqrt{nM} < j \le n$. In particular, we can assume that $M < n$. If $1 - 2/n \le t \le 1 + 2/n$, then $\frac{1}{|1-t|} \ge \frac{n}{2}$ and so, (6.6) follows by a similar argument to (6.7). Thus, we can further assume that $t < 1 - \frac{2}{n}$.

For each $\sqrt{nM} < j \le n$, we have from (2.5)

$$\sum_{i=0}^{\lfloor j-\sqrt{nM}\rfloor} (j-i)^2 c_i^2 c_j^2 t^{2i+2j} = O\left(j^2 c_j^2 c_{\lfloor j-\sqrt{nM}\rfloor}^2 t^{2j} \sum_{i=0}^{\infty} t^{2i}\right) = O\left(n^{2\rho+2} c_{\lfloor j-\sqrt{nM}\rfloor}^2 M t^{2j}\right). \tag{6.8}$$

We will now show that

$$n^{2\rho+2} c_{\lfloor j-\sqrt{nM}\rfloor}^2 M t^{2j} = \frac{O(n)}{1-t} \mathcal{P}(t) c_{\lfloor j-\sqrt{nM}\rfloor}^2 t^{2\lfloor j-\sqrt{nM}\rfloor}, \tag{6.9}$$

which is equivalent to $n^{2\rho+1} = O(1)\mathcal{P}(t)(1-\frac{1}{M})^{-2\sqrt{nM}}$ for some constant C_3. This is true because the right-hand side is at least $(\sum_{i=\lceil M/2\rceil}^{M} c_i^2 (1-\frac{1}{M})^{2i}) e^{2\sqrt{n/M}} \gg M^{2\rho+1} e^{2\sqrt{n/M}} \gg n^{2\rho+1}$; note that we assumed that $M = \frac{1}{1-t} < n$.

From (6.7), (6.8) and (6.9), we obtain (6.6).

The proof for (6.4) follows nearly the same lines with

$$\mathcal{I}_{\widetilde{Q}} = \mathbf{E} \sum_{i=1}^{n} \mathbf{1}_{\pm[10^{-7}\delta, 10^{-7}\delta+10^{-3}\delta^{\alpha+1}]}\big(\zeta_i^{\widetilde{Q}} - x_0\big)$$

and as in the proof of Lemma 5.1, the estimates for $\widetilde{Q}$ will be similar to the case $\rho = 0$ for $\widetilde{P}$. In particular, the handling of the corresponding second integral of (6.5) is similar (exploiting ingredients from the proof of Lemma 5.1), and for the first integral we may upper bound it by that of Kac polynomials by comparing c_j with the hyperbolic coefficients and using Lemma 10.6. This completes the proof. □

7. Proof of complex local universality for series. PROOF OF THEOREMS 2.10 AND 2.13. First, let us make some observations about the series P_{PS} under Condition 1:

1. P_{PS} converges uniformly in every compact set in $\mathbf{D}$ a.e.

Indeed, let $\Omega_n = \{\omega : |\xi_i(\omega)| \le n + i^n, \forall i \ge 0\}$. Then $\Omega_1 \subset \Omega_2 \cdots \subset \Omega_n \cdots$, and $\Omega = \bigcup_{n=1}^{\infty} \Omega_n$. On each Ω_n, P_{PS} converges uniformly on compact sets in $\mathbf{D}$.

Moreover, P_{PS} does not extend analytically to any domain larger than the unit disk (see, for instance, [13], Lemma 2.3.3).

2. By the Lebesgue's dominated convergence theorem, $\mathbf{Var}(P_{\mathrm{PS}}(z)) = \sum_{n=0}^{\infty} |c_n|^2 |z|^{2n}$.

3. For every $0 < \delta \le 1 - \frac{1}{C}$, $z \in A(0, 1-2\delta, 1-\delta]$, $k \ge 1$, one has

$$\mathbf{E}\big[N_{P_{\mathrm{PS}}}\big(B(z, \delta/10)\big)\big]^k < \infty.$$

This follows from Proposition 4.6 by setting $\lambda = 2^n$ with $n = 1, 2, 3, \ldots$ and shows that the integrals in the statements of Theorems 2.10 and 2.13 are well defined.

Now the proofs of Theorems 2.10 and 2.13 for any $0 < \delta \le \frac{1}{C}$ follow exactly the same lines as the proofs of Theorems 2.3 and 2.4 for the case $\frac{\log^2 n}{n} \le \delta \le \frac{1}{C}$ with the n in the latter proofs being replaced by ∞. □

PROOF OF COROLLARY 2.11. The corollary follows from Theorem 2.3 with the two sequences of random variables (ξ_n) and $(\xi_n e^{\sqrt{-1}n\theta})$. □

PROOF OF COROLLARY 2.12. Observe that by the change of variables formula, with respect to the rescaling formula 2.7, one has

$$\rho_{\check{P}}^{(k)}(w_1, \dots, w_k) = (10^{-3}\delta)^{2k} \rho_P^{(k)}(10^{-3}\delta w_1, \dots, 10^{-3}\delta w_k). \tag{7.1}$$

Let $\widetilde{P}_{\mathrm{PS}}$ be the hyperbolic power series with ξ's being i.i.d. standard complex Gaussian. By Theorem 2.10, we have

$$\begin{aligned} &\biggl|\int_{\mathbb{C}^k} G(w)(10^{-3}\delta_0)^{2k} \rho_{P_{\mathrm{PS}}}^{(k)}(z + 10^{-3}\delta_0 w)\, dw_1 \cdots dw_k \\ &\quad - \int_{\mathbb{C}^k} G(w)(10^{-3}\delta_0)^{2k} \rho_{\widetilde{P}_{\mathrm{PS}}}^{(k)}(z + 10^{-3}\delta_0 w)\, dw_1 \cdots dw_k\biggr| \le C'\delta_0^c. \end{aligned} \tag{7.2}$$

As proven in Proposition 2.3.4 in [13], the zero set of $\widetilde{P}_{\mathrm{PS}}$ is invariant in distribution under the transformations ϕ. Thus,

$$\begin{aligned} &\int_{\mathbb{C}^k} G(w)(10^{-3}\delta_0)^{2k} \rho_{\widetilde{P}_{\mathrm{PS}}}^{(k)}(z + 10^{-3}\delta_0 w)\, dw_1 \cdots dw_k \\ &\quad = \int_{\mathbb{C}^k} H(w)(10^{-3}\delta_1)^{2k} \rho_{\widetilde{P}_{\mathrm{PS}}}^{(k)}(t + 10^{-3}\delta_1 w)\, dw_1 \cdots dw_k. \end{aligned} \tag{7.3}$$

Thus, it remains to show that

$$\begin{aligned} &\biggl|\int_{\mathbb{C}^k} H(w)(10^{-3}\delta_1)^{2k} \rho_{P_{\mathrm{PS}}}^{(k)}(t + 10^{-3}\delta_1 w)\, dw_1 \cdots dw_k \\ &\quad - \int_{\mathbb{C}^k} H(w)(10^{-3}\delta_1)^{2k} \rho_{\widetilde{P}_{\mathrm{PS}}}^{(k)}(t + 10^{-3}\delta_1 w)\, dw_1 \cdots dw_k\biggr| \le C'\delta_1^c. \end{aligned} \tag{7.4}$$

Recall that the hyperbolic area is defined by $\mathrm{Area}(B) := \int_B \frac{dm(z)}{(1-|z|^2)^2}$ for every Borel set $B \subset \mathbf{D}$. By the change of variables formula, one can prove that if ϕ is a hyperbolic transformation then ϕ preserves the hyperbolic area, that is, $\mathrm{Area}(B) = \mathrm{Area}(\phi(B))$. Moreover, ϕ maps circles in $\mathbf{D}$ into circles in $\mathbf{D}$ (see, for instance, [34], Section 14.3).

Now, since ϕ maps z_j to t_j with $|z_j| \in [1 - 2\delta_0, 1 - \delta_0]$ and $t_j \in [1 - 2\delta_1, 1 - \delta_1]$, one has

$$\phi(\mathbf{D}(z_j, \delta_0/s)) \subset \mathbf{D}(t_j, 10\delta_1/s) \tag{7.5}$$

for every $s \geq 25$. Indeed, assume that $t_j \in \phi(\mathbf{D}(z_j, \delta_0/s)) = \mathbf{D}(t, r)$. Then $\text{Area}(\mathbf{D}(z_j, \delta_0/s)) = \text{Area}(\mathbf{D}(t, r))$. We have

$$\text{Area}(\mathbf{D}(z_j, \delta_0/s)) = \int_{\mathbf{D}(z_j,\delta_0/s)} \frac{dm(z)}{(1-|z|^2)^2} \leq \frac{\pi\delta_0^2}{s^2} \frac{1}{(\delta_0 - \delta_0/s)^2} \leq \frac{\pi}{(s-1)^2}.$$

The radius r cannot be larger than $1/3$ because otherwise, there exits some t' between t and t_j such that $|t' - t_j| = \delta_1/2$. And so $\mathbf{D}(t', \delta_1/2) \subset \mathbf{D}(t, r)$, but then

$$\text{Area}(\mathbf{D}(t', \delta_1/2)) \geq \frac{\pi\delta_1^2}{4} \frac{1}{(1-(1-3\delta_1)^2)^2} \geq \frac{\pi}{144} > \phi(\mathbf{D}(z_j, \delta_0/s)) \tag{7.6}$$

which is impossible. So, $r \leq 1/3$, and hence, for every $z \in \mathbf{D}(t, r)$, $|z| \geq |t_j| - 2r \geq 1 - 2\delta_1 - 2r > \frac{1}{3} - 2\delta_1 > 0$. Therefore, $\text{Area}(\mathbf{D}(t, r)) \geq \frac{\pi r^2}{16(\delta_1 + r)^2}$. Comparing this with (7.5), we conclude that $r \leq \frac{4\delta_1}{s-5}$. Hence, $\mathbf{D}(t, r) \subset \mathbf{D}(t_j, \frac{8\delta_1}{s-5}) \subset \mathbf{D}(t_j, \frac{10\delta_1}{s})$, proving (7.5).

From this and the assumption that G is supported in $B(0, 10^{-4})^k$, one can deduce that H is supported in $B(0, 10^{-3})^k$. The inequality (7.4) will then follow from Theorem 2.3 if we can show that $|\nabla^a H(z)| \leq C$ for all $0 \leq a \leq 2k + 4$ and $z \in \mathbb{C}^k$, which in turn follows from the bounds:

$$|(\phi^{-1})^{(n)}(z)| \leq C_n \frac{\delta_0}{\delta_1^n} \qquad \forall n \geq 0, \forall z \in \mathbf{D}(t_i, 10^{-6}\delta_1), \tag{7.7}$$

where C_n is a constant depending on n.

Hence, it remains to show (7.7). Since $\phi^{-1}(t_j) = z_j$, there exists some $\theta \in [0, 2\pi)$ such that $\phi^{-1}(z) = \varphi_{-z_j}(e^{\sqrt{-1}\theta}\varphi_{t_j}(z))$ for all $z \in \mathbf{D}$ where $\varphi_\alpha = \frac{z-\alpha}{1-z\bar{\alpha}}$ (see, for instance, [34], Sections 12.4, 12.5). Since $e^{\sqrt{-1}\theta}$ does not change the magnitudes of the derivatives, we can assume without loss of generality that $\theta = 0$. Now, by direct computation, we have

$$\begin{aligned} |\varphi_{t_j}^{(m)}(z)| &= \left| \frac{m!(1-|t_j|^2)\bar{t}_j^{m-1}}{(1-\bar{t}_j z)^{m+1}} \right| \\ &\leq \frac{C_m \delta_1}{\delta_1^{m+1}} = \frac{C_m}{\delta_1^m} \qquad \forall m \geq 0, \forall z \in \mathbf{D}(t_i, 10^{-6}\delta_1). \end{aligned} \tag{7.8}$$

For $z \in \mathbf{D}(t_i, 10^{-6}\delta_1)$, set $w = \varphi_{t_j}(z) = \frac{z-t_j}{1-\bar{t}_j z} \in \mathbf{D}(0, 10^{-5})$. And

$$|\varphi_{-z_j}^{(m)}(w)| = \left| \frac{m!(1-|z_j|^2)\bar{z}_j^{m-1}}{(1-\bar{z}_j w)^{m+1}} \right| \leq C_m \delta_0 \qquad \forall m \geq 0, \forall w \in \mathbf{D}(0, 10^{-5}). \tag{7.9}$$

Combining (7.8) and (7.9), we obtain (7.7) and complete the proof. $\square$

8. Proof of Theorem 2.8, part I: Reduction to the case $M = m = 1$. We begin the proof of Theorem 2.8 in this section. In this section, ξ_k's are i.i.d. normalized Gaussian and $(c_k)_{k\ge 0}$ is a sequence of deterministic real numbers satisfying the following assumptions. For some $N_0 \ge 0$ and $0 < m \le M < \infty$, it holds that

$$m\sqrt{h(k)} \le |c_k| \le M\sqrt{h(k)}, \qquad N_0 \le k \le n, \qquad \max_{0\le k<N_0} c_k^2 \le C_1 M.$$

Below, we let $N_n(I)$ be the number of real zeros of $P_n(t) = \sum_{k=0}^n c_k \xi_k t^k$ that are inside I for any $I \subset \mathbb{R}$. For brevity, we will sometimes write $N_n(a,b) = N_n((a,b))$, $N_n[a,b] = N_n([a,b])$, etc.

Our main goal of this section is to reduce the theorem to the simpler case $M = m = 1$.

By Edelman–Kostlan [10], the density function for the distribution of the real zeros for $P_n(x)$ is $\rho_n(t) = \frac{1}{\pi}\|\gamma_n'(t)\|$ where $\gamma_n(t)$ is the unit vector in the direction of $v_n(t) := (c_0, c_1 t, \dots, c_n t^n)$. It was shown in [10] that

$$\|\gamma_n'(t)\|^2 = \left(\frac{\|v_n'(t)\|}{\|v_n(t)\|}\right)^2 - \left(\frac{v_n(t)\cdot v_n'(t)}{\|v_n(t)\|^2}\right)^2. \tag{8.1}$$

From (8.1), it follows that ρ_n is an even function of t.

By elementary computation, for any $n \ge 0$ and any sequence (x_k) and (y_k), we have

$$\left(\sum_{k=0}^n x_k^2\right)\left(\sum_{k=0}^n y_k^2\right) - \left(\sum_{k=0}^n x_k y_k\right)^2 = \sum_{k,m}(x_k y_m - x_m y_k)^2.$$

It follows that

$$\begin{aligned}\rho_n(t)^2 &= \frac{1}{\pi^2}\frac{\|v_n'(t)\|^2\|v_n(t)\|^2 - [v_n'(t)\cdot v_n(t)]^2}{\|v_n(t)\|^4}\\ &= \frac{1}{\pi^2}\frac{\sum_{0\le k,m\le n}(m-k)^2 c_k^2 c_m^2 t^{2m+2k-2}}{(\sum_{k=0}^n c_k^2 t^{2k})^2}.\end{aligned}$$

Thus, for $|t|$ comparable to 1 we have $\rho(t) = O(n)$, therefore, $\mathbf{E}N_n(1-\frac{c}{n}, 1+\frac{c}{n}) = O(1)$ for any absolute constant $c > 0$. Furthermore, by scaling invariant one sees that

COROLLARY 8.1. *Suppose that for $0 < m \le M < \infty$ we have $m|b_k| \le |a_k| \le M|b_k|$ for every $k = 0, \dots, n$. Let N_n and $\widetilde{N}_n$ respectively count the real zeros of random polynomials associated with $a_0, \dots, a_n$ and $b_0, \dots, b_n$. Then*

$$\frac{m^2}{M^2}\mathbf{E}\widetilde{N}_n \le \mathbf{E}N_n \le \frac{M^2}{m^2}\mathbf{E}\widetilde{N}_n.$$

Thanks to Corollary 8.1, it suffices to prove Theorem 2.8 for $m = M = 1$. We will free the symbols m and M so that they could be used for unrelated purposes later.

We now describe the high-level overview of the rest of the proof of Theorem 2.8. Thanks to Lemma 2.5, it remains to count the number of real zeros near the critical points $x = -1$ and $x = 1$. By symmetry, it suffices to consider a small neighborhood of 1, which we will discuss in the next two section: Section 9 will discuss estimates for the density function near 1 and Section 10 will use these results to estimate the average number of real zeros near 1.

9. Proof of Theorem 2.8, part II: Estimates for the density function near ± 1. In this section, we prove some estimates for ρ_n near ± 1.

Below, for $x \geq 0$ let $f_n(x) = \sum_{0 \leq k \leq n} c_n^2 x^n$, clearly $\mathbf{Var}[P_n(t)] = f_n(t^2)$ so our notational convention is to think of x as t^2.

Our general framework for the analysis in this section will be under the heuristics that $f_n(x)$ converges fairly rapidly to some $f_\infty(x)$ as $n \to \infty$. This convergence essentially leads to the convergence of ρ_n to some limit ρ_∞. The local average number of real zeros of P_n is essentially decided by the local behavior of ρ_∞ and the rate of the convergence $f_n \to f_\infty$. For instance, if $P_n(t) = \sum_{k=0}^n c_k \xi_k t^k$ where ξ_k are i.i.d. normalized Gaussian and c_k are independent of n then the natural choice for f_∞ would be $f_\infty(x) = \sum_{k=0}^\infty c_k^2 x^k$, and the convergence $f_n \to f_\infty$ holds for x inside the radius of convergence of f_∞. On the other hand, our approach is applicable even if c_k depend on n, and does not require the polynomially growing assumptions on c_k.

To motivate the definition of ρ_∞, we let $g_n(x) = \log f_n(x)$, and note the following.

LEMMA 9.1. *For every n, it holds that*

$$\rho_n(t) = \frac{1}{\pi}\bigl(g_n'(t^2) + t^2 g_n''(t^2)\bigr)^{1/2}. \tag{9.1}$$

PROOF. Let $v_n(t)$ denote the vector $(c_0, c_1 t, \dots, c_n t^n)$. Clearly,

$$\|v_n(t)\|^2 = \sum_{0 \leq k \leq n} c_k^2 t^{2k} = f_n(t^2),$$

$$v_n'(t) \cdot v_n(t) = \sum_{0 \leq k \leq n} k c_k^2 t^{2k-1} = \frac{1}{2}\frac{d}{dt}\bigl(\|v_n(t)\|^2\bigr) = t f_n'(t^2),$$

$$\|v_n'(t)\|^2 = \sum_{0 \leq k \leq n} k^2 c_k^2 t^{2k-2} = \frac{1}{4t}\frac{d}{dt}\left(t\frac{d}{dt}\bigl(\|v_n(t)\|^2\bigr)\right) = f_n'(t^2) + t^2 f_n''(t^2).$$

The desired claim now follows from the Edelman–Kostlan formula (8.1):

$$\pi^2\rho_n(t)^2 = \left(\frac{\|v_n'(t)\|}{\|v_n(t)\|}\right)^2 - \left(\frac{v_n(t)\cdot v_n'(t)}{\|v_n(t)\|^2}\right)^2$$
$$= \frac{f_n'(t^2)}{f_n(t^2)} + t^2\frac{f_n''(t^2)f_n(t^2) - [f_n'(t^2)]^2}{[f_n(t^2)]^2}$$
$$= g_n'(t^2) + t^2 g_n''(t^2). \qquad \square$$

Let $0 \le \beta < \alpha < \infty$ such that for $x \in (\alpha^2, \beta^2)$ the limit $f_\infty(x) := \lim_{n\to\infty} f_n(x)$ exists and is continuously twice differentiable on this interval. Let $g_\infty(x) = \log f_\infty(x)$ and define

$$\rho_\infty(t) := \frac{1}{\pi}\sqrt{g_\infty'(t^2) + t^2 g_\infty''(t^2)}. \tag{9.2}$$

Motivated by Lemma 9.1, under some mild assumptions one expects that $\rho_n(t)$ converges to $\rho_\infty(t)$ for $\beta < |t| < \alpha$. The precise estimates will be discussed below.

Note that the current analysis is only directly applicable to count the number of real zeros inside $(-\alpha,\alpha)$ near $\pm\alpha$. For $\mathbb{R}\setminus(-\alpha,\alpha)$, we will pass to the reciprocal polynomial $\widetilde{P}_n(t) = \frac{1}{c_n}t^n P_n(\frac{1}{t})$ and apply the argument to $\widetilde{P}_n$, which is also a Gaussian random polynomial.

9.1. *Convergence of ρ_n.*

THEOREM 9.2. *Let $u_n(x) := \frac{f_n(x)}{f_\infty(x)}$. Assume that $I_n \subset (\beta,\alpha)$ is an interval (whose endpoints may depend on n) such that $u_n(t^2) \ge c_0$ for $|t| \in I_n$ for some fixed constant $c_0 > 0$. Then uniformly over $\{|t| \in I_n\}$ it holds that*

$$\rho_n(t) = \rho_\infty(t) + O\big(|u_n'(t^2)|^{1/2} + |u_n'(t^2)| + |u_n''(t^2)|^{1/2}\big).$$

PROOF. Let $D_n(x) = \log f_n(x) - \log f_\infty(x)$. Using Lemma 9.1, we have

$$|\rho_n(t) - \rho_\infty(t)| \le |\rho_n(t)^2 - \rho_\infty(t)^2|^{1/2} \le |D_n'(t^2)|^{1/2} + \alpha|D_n''(t^2)|^{1/2}.$$

On the other hand, let $x = t^2$ where $t \in I_n$, then $u_n(x) \ge c_0 > 0$, therefore,

$$D_n'(x) = \frac{u_n'(x)}{u_n(x)} = O\big(u_n'(x)\big)$$
$$D_n''(x) = \frac{u_n''(x)}{u_n(x)} - \left(\frac{u_n'(x)}{u_n(x)}\right)^2 = O\big(D_n'(x)^2 + |u_n''(x)|\big)$$

and the desired estimate immediately follows. $\square$

We remark that the assumption $u_n(t^2) \ge c_0 > 0$ uniformly over $|t| \in I_n$ in Theorem 9.2 is fairly mild, since one has $u_n(x) \to 1$ as $n \to \infty$.

9.2. *Blowup nature of* ρ_∞. It follows from Theorem 9.2 that the leading asymptotics of ρ_n on I_n is determined by two factors: the size of $u_n = f_n/f_\infty$ (and its first two derivatives), and the possible blowup of ρ_∞, which typically could happen near the endpoint of I_n. By (9.2) depends on the blowup nature of f_∞ there. For the polynomially growing setting of Theorem 2.8 (and with the normalization $M = m = 1$), one expects that f_∞ blows up polynomially near the endpoints of its convergence interval. This will lead to a simple pole for ρ_∞, as proved in the following lemma.

LEMMA 9.3. *Let* $0 \le \beta < \alpha < \infty$ *and* $\gamma \ge 0$. *Assume that* $\log f_\infty(x) + \gamma \log|x - \alpha^2|$ *has two uniformly bounded derivatives for* $x \in (\beta^2, \alpha^2)$. *Then the following holds uniformly over* $|t| \in (\beta, \alpha)$:

$$\rho_\infty(t) = \frac{\alpha\sqrt{\gamma}}{\pi|t^2 - \alpha^2|} + O(1).$$

PROOF. Recall that $g_\infty = \log f_\infty$. For $|t| \in (\beta, \alpha)$, by the given assumption we have

$$g'_\infty(t^2) = -\frac{\gamma}{t^2 - \alpha^2} + O(1), \qquad g''_\infty(t^2) = \frac{\gamma}{(t^2 - \alpha^2)^2} + O(1).$$

Using (9.2), we obtain

$$\begin{aligned}\rho_\infty(t)^2 &= \frac{1}{\pi^2}\big(g'_\infty(t^2) + t^2 g''_\infty(t^2)\big) = \frac{1}{\pi^2}\left(-\frac{\gamma}{t^2 - \alpha^2} + \frac{t^2\gamma}{(t^2 - \alpha^2)^2}\right) + O(1)\\ &= \frac{1}{\pi^2}\frac{\gamma\alpha^2}{(t^2 - \alpha^2)^2} + O(1).\end{aligned}$$

Since $\rho_\infty \ge 0$, the desired conclusion follows immediately. □

10. Proof of Theorem 2.8, part III: Counting real zeros near ±1. Recall that $h(k) = \sum_{j=0}^{d} \alpha_j L_j(L_j + 1)\cdots(L_j + k - 1)/k!$ with nonzero coefficients, and for some fixed $N_0 \ge 0$ the following hold:

- for every $N_0 \le k \le n$ it holds that $|c_k| = \sqrt{h(k)}$.
- for some C_1 fixed we have $\max_{0 \le k < N_0} |c_k| < C_1$.

Without loss of generality, assume that $\alpha_d = 1$.

To count the real zeros near ±1 of $P_n(t) = \sum_{k=0}^{d} c_k \xi_k t^k$, we separate the treatment of the inside and outside into two results, Lemmas 10.1 and 10.2 below. In the following two results, the implicit constants may depend on N_0, C_1 and h.

LEMMA 10.1. *For some* $\beta \in (0, 1)$ *that depends only on* h, N_0, C_1, *it holds that*

$$\mathbf{E}N_n(\{\beta \le |t| \le 1\}) = \frac{\sqrt{\deg(h) + 1}}{\pi}\log n + O(1).$$

LEMMA 10.2. *It holds that*

$$\mathbf{E}N_n\big([-2,-1]\cup[1,2]\big)=\frac{\log n}{\pi}+O(1).$$

REMARK. It is clear that Theorem 2.8 follows from Lemma 2.5, Lemma 10.1, Lemma 10.2. Therefore this section completes the proof of Theorem 2.8.

For convenience of notation, in the rest of the section let

$$g(x)=\sum_{k=0}^{N_0-1}[c_k^2-h(k)]x^k. \tag{10.1}$$

It follows that $f_n(x)=g(x)+\sum_{k=0}^{n}h(k)x^k$. Furthermore, $g(x)$ and its derivatives are uniformly bounded on any compact subset of $\mathbb{R}$ with bounds depending on C_1 and N_0 and h. This fact will be used implicitly below.

For any $L\in\mathbb{R}$, we also let

$$f_{n,L}(x)=\sum_{k=0}^{n}b_{k,L}x^k,\qquad b_{k,L_d}:=L_d\cdots(L_d+k-1)/k!.$$

10.1. *Proof of Lemma* 10.1. Clearly, $f_n(x)\to f_\infty(x)$ for $|x|<1$, and $f_\infty(x)=g(x)+\sum_{k=0}^{\infty}h(k)x^k$.

Using the binomial expansion of $(1-x)^{-L}$, we obtain

$$f_\infty(x)=g(x)+\sum_{m=0}^{d}\alpha_m(1-x)^{-L_m}$$

for every $x\in[-1,1)$. Since $\alpha_d>0$ and $L_d>\cdots>L_0>0$, it follows that

$$\log f_\infty(x)+L_d\log(1-x)$$

is bounded uniformly over $x\in[\beta^2,1)$ for some $\beta\in(0,1)$ depending only on N_0, C_1 and h. We furthermore choose $\beta\in(0,1)$ to be sufficiently close to 1 such that $|(\frac{d}{dx})^j f_\infty(x)|\approx(1-x)^{-L_d-j}$[5] uniformly over $x\in[\beta^2,1)$ where $j=0,1,2$. Now, for $c_0=\min(L_d,\min_j(L_j-L_{j-1}))>0$, it is clear that the jth derivative of $\log f_\infty(x)+L_d\log(1-x)=\log[(1-x)^{L_d}f_\infty(x)]$ is bounded above by $O((1-x)^{c_0-j})$. Using (9.2) and argue as in the proof of Lemma 9.3, it follows that

$$\rho_\infty(t)=\frac{\sqrt{L_d}}{\pi(1-t^2)}+O\big((1-t^2)^{\frac{c_0}{2}-1}\big)$$

uniformly over $|t|\in[\beta,1)$. Now recall that $L_d\equiv\deg(h)+1$. By the symmetry of the real zeros distribution, we have

$$\mathbf{E}N_n\big(\{\beta\le|t|\le1\}\big)=2\int_\beta^{1-\frac{c}{n}}\rho_n(t)\,dt+O(1), \tag{10.2}$$

[5] We say that $f\approx g$ if there exist constants c, C such that $cf\le g\le Cf$.

where $c > 0$ is any fixed constant. We will use the above estimate for ρ_∞ to show the following.

LEMMA 10.3. *For some fixed c sufficiently large, it holds uniformly over $|t| \in (\beta, 1 - \frac{c}{n})$ that*

$$\rho_n(t) = \frac{\sqrt{\deg(h)+1}}{2\pi(1-|t|)} + O(1) + O\big((1-|t|)^{\frac{c_0}{2}-1}\big) + O\left(\frac{[n(1-t^2)]^{(L_d+1)/2}|t|^n + [n(1-t^2)]^{L_d}|t|^{2n}}{1-|t|}\right).$$

We first show that this lemma implies the desired estimate for Lemma 10.1. Indeed, notice that for every $\alpha > 0$ we have $\frac{\alpha\cdots(\alpha+k-1)}{k!} \approx k^{\alpha-1}$ and $\sum_{k=1}^n k^{\alpha-1} \approx n^\alpha$, therefore, we obtain the following uniform estimates (over $0 \le x \le 1$):

$$n^\alpha x^n \le C \sum_{k=0}^n \frac{\alpha\cdots(\alpha+k-1)}{k!} x^k \le \frac{C}{(1-x)^\alpha}.$$

Combining this with Lemma 10.3, we obtain the uniform estimate:

$$\rho_n(t) = \frac{\sqrt{\deg(h)+1}}{2\pi(1-t)} + O\big((1-|t|)^{\frac{c_0}{2}-1}\big) + O\left(\frac{1}{n(1-|t|)^2}\right)$$

over $t \in [\beta, 1 - \frac{c}{n}]$ where $c > 0$ is any fixed large constant. Together with (10.2), we obtain

$$\mathbf{E}N_n\big(\{\beta \le |t| \le 1\}\big) = \frac{\sqrt{\deg(h)+1}}{\pi}\log n + O(1),$$

as stated in Lemma 10.1.

We now prove Lemma 10.3. The proof of this lemma relies on the following estimates for f_n.

LEMMA 10.4. *For each $j = 0, 1, 2$, it holds uniformly over $x \in [\beta^2, 1)$ that*

$$\left(\frac{d}{dx}\right)^j \big(f_n(x) - f_\infty(x)\big) = O\left(\frac{(1 + [n(1-x)]^{L_d+j-1})x^{n+1}}{(1-x)^{L_d+j}}\right)$$

and it holds uniformly over $x \in [-1, 0]$ that $f_n(x) = O((1+x)^{-(L_d-1)})$.

We first prove Lemma 10.3 using Lemma 10.4. Let $u_n = f_n(x)/f_\infty(x)$. By Lemma 10.4, uniformly over $x \in (\beta^2, 1)$ and $j = 0, 1, 2$ it holds that

$$\left(\frac{d}{dx}\right)^j \big(u_n(x) - 1\big) = O\big((1-x)^{-j}\big(1 + [n(1-x)]^{L_d+j-1}\big)x^n\big).$$

In particular, for c large and $\beta^2 \le x \le 1 - \frac{c}{n}$ we have $u_n(x) = 1 + O(x^{n/2}) = 1 + O(e^{-c/2})$, and thus, $u_n(x) \in [\frac{1}{2}, \frac{3}{2}]$. Therefore, Theorem 9.2 is applicable, and we obtain the desired estimate of Lemma 10.3.

PROOF OF LEMMA 10.4. Consider $x \in [\beta^2, 1)$. It suffices to show that for every $L > 0$ and each $0 \le j \le 2$ the following holds uniformly:

$$\left(\frac{d}{dx}\right)^j \left(-\frac{1}{(1-x)^L} + f_{n,L}(x)\right) = O_L\left(\frac{(1+[n(1-x)]^{L+j-1})x^{n+1}}{(1-x)^{L+j}}\right). \tag{10.3}$$

Similarly, for $x \in [-1, 0]$ it suffices to show that for any $L > 0$,

$$f_{n,L}(x) = O\big((1+x)^{L-1}\big). \tag{10.4}$$

Observe that

$$\frac{d}{dx}\left(-\frac{1}{(1-x)^L} + f_{n,L}(x)\right) = L\left(-\frac{1}{(1-x)^{L+1}} + f_{n-1,L+1}(x)\right)$$

therefore, in (10.3) we may assume that $j = 0$.

Now for $0 \le x < 1$ we have

$$\begin{aligned} -\frac{1}{(1-x)^L} &+ \sum_{k=0}^{n} \frac{L\cdots(L+k-1)}{k!} x^k \\ &= \sum_{k=n+1}^{\infty} \frac{L\cdots(L+k-1)}{k!} x^k \\ &= x^{n+1} \sum_{k=0}^{\infty} \frac{L\cdots(L+k+n)}{(n+1+k)!} x^k. \end{aligned} \tag{10.5}$$

Now we will use the standard asymptotic estimate for generalized binomial coefficients

$$\frac{L(L+1)\cdots(L+k-1)}{k!} \approx C k^{L-1}$$

as $k \to \infty$ where C depends on L. It follows that

$$\begin{aligned} \frac{L(L+1)\cdots(L+k+n)}{(n+1+k)!} &\le C \frac{L(L+1)\cdots(L+k-1)}{k!} \left(\frac{n+k+1}{k}\right)^{L-1} \\ &\le C \frac{L(L+1)\cdots(L+k-1)}{k!} \left(1 + \frac{(n+1)^{L-1}}{k^{L-1}}\right) \\ &\le C \frac{L(L+1)\cdots(L+k-1)}{k!} + (n+1)^{L-1} \end{aligned}$$

(in the last estimate we use the asymptotic for generalized binomial coefficients again). Using (10.5) and the binomial expansion, it follows that

$$
\begin{aligned}
&-\frac{1}{(1-x)^L}+\sum_{k=0}^{n}\frac{L(L+1)\cdots(L+k-1)}{k!}x^k\\
&\quad\le x^{n+1}\bigl[(1-x)^{-L}+(n+1)^{L-1}(1-x)^{-1}\bigr]\\
&\quad\le C\bigl(\bigl(1+\bigl[n(1-x)\bigr]^{L-1}\bigr)(1-x)^{-L}x^{n+1}\bigr)
\end{aligned}
$$

giving (10.3).

For $x\in[-1,0]$, we will use the following recursive formulas.

LEMMA 10.5. *For any $x\neq 1$, it holds that*

$$
f_{n,L}(x)=\frac{f_{n,L-1}(x)}{1-x}-\frac{L\cdots(L+n-1)}{n!}\frac{x^{n+1}}{1-x}.
$$

PROOF. We have

$$
\begin{aligned}
f_{n,L}(x)&=1+Lx+\frac{L(L+1)}{2}x^2+\cdots+\frac{L(L+1)\cdots(L+n-1)}{n!}x^n,\\
xf_{n,L}(x)&=x+Lx^2+\cdots+\frac{L\cdots(L+n-1)}{n!}x^{n+1},\\
(1-x)f_{n,L}(x)&=1+(L-1)x+\cdots+\frac{L(L+1)\cdots(L+n-2)(L-1)}{n!}x^n\\
&\quad-\frac{L\cdots(L+n-1)}{n!}x^{n+1},\\
&=f_{n,L-1}(x)-\frac{L\cdots(L+n-1)}{n!}x^{n+1}
\end{aligned}
$$

and the desired claim follows. □

For $x\in[-1,0]$ it is clear that $\frac{L\cdots(L+n-1)}{n!}\frac{x^{n+1}}{1-x}=O(n^{L-1}|x|^n)=O(\frac{1}{(1+x)^{L-1}})$. Thus, without loss of generality we may assume that $0<L\le 1$. For this L, for $x\in[-1,0]$ it is clear that $f_{n,L}$ is an alternating sum whose terms have decreasing modulus, and could be easily bounded by $O(1)$ uniformly over $x\in[-1,0]$. □

10.2. *Proof of Lemma* 10.2. Thanks to the symmetry of the distribution of the real zeros, we have

$$
\mathbf{E}N_n(\{1\le|t|\le 2\})=2\mathbf{E}\widetilde{N}_n\left(\frac{1}{2},1\right)=2\int_{\frac{1}{2}}^{1-\frac{c}{n}}\tilde{\rho}_n(t)\,dt+O(1),\tag{10.6}
$$

where $\widetilde{N}_n$ and $\tilde{\rho}_n$ are respectively the number of real zeros and the density of the real zeros distribution for the normalized reciprocal polynomial

$$\widetilde{P}_n(t) = \sum_{k=0}^{n} \frac{c_{n-k}}{c_n} \xi_k t^k.$$

We note that $|c_n| = \sqrt{h(n)}$ so $c_n \neq 0$ for n sufficiently large, so $\widetilde{P}_n$ is well defined.

Let $\tilde{f}_n(x)$ denote the corresponding variance function

$$\tilde{f}_n(x) = \sum_{k=0}^{n} \frac{c_{n-k}^2}{c_n^2} x^k \equiv \frac{x^n f_n(1/x)}{c_n^2}.$$

As we will see, for any $0 \le x < 1$ the sequence $\tilde{f}_n(x)$ converges to $\tilde{f}_\infty(x) := \frac{1}{1-x}$, which suggests that $\tilde{\rho}_n(t)$ is asymptotically $\frac{1}{2\pi(1-t)}$ for $t \in [\frac{1}{2}, 1)$. In fact, we will show the following.

LEMMA 10.6. *Suppose that $c > 0$ is a sufficiently large fixed constant. Then uniformly over $t \in [\frac{1}{2}, 1 - \frac{c}{n}]$, it holds that*

$$\tilde{\rho}_n(t) = \frac{1}{2\pi(1-t)}\bigl(1 + O\bigl(n^{L_{d-1}-L_d}\bigr)\bigr) + O(1) + O\biggl(\frac{1}{n(1-t)^2} + \frac{1}{\sqrt{n(1-t)^3}}\biggr).$$

From the following computation, Lemma 10.6 and (10.6) imply the desired estimate for Lemma 10.2:

$$\begin{aligned}\mathbf{E}N_n\bigl(\{1 \le |t| \le 2\}\bigr) &= 2\int_{\frac{1}{2}}^{1-\frac{c}{n}} \frac{1}{2\pi(1-t)}\,dt + O(1) \\ &\quad + O\biggl(\int_{\frac{1}{2}}^{1-\frac{c}{n}} \frac{1}{n(1-t)^2} + \frac{1}{n^{1/2}(1-t)^{3/2}}\,dt\biggr) \\ &= \frac{\log n}{\pi} + O(1).\end{aligned}$$

To prove Lemma 10.6, we reduce the problem to the hyperbolic setting. As we will see, $\tilde{f}_n(x)$ converges to $\tilde{f}_\infty(x) = \frac{1}{1-x}$ for every $x \in [0, 1)$ sufficiently close to 1, say $x \in [1/2, 1)$. Our proof will make use of the density comparison results developed in the previous section, Theorem 9.2 and Lemma 9.3, relying on various estimates for $\tilde{f}_n(x)/\tilde{f}_\infty(x)$ and its first two derivatives. It is clear that modulo the contribution of g [defined in (10.1)] which will be shown to be very small, $\tilde{f}_n(x)$ is a linear combination of $\tilde{f}_{n,L_j}$ where the linear coefficient for $\tilde{f}_{n,L_d}$ is $1 + O(n^{-c})$ and the linear coefficients of other terms are $O(n^{-c})$ where $c = L_d - L_{d-1}$. Thus, it suffices to consider the setting when $f_n = g + f_{n,L_d}$, which we assume below.

We first establish some basic estimates for $f_{n,L}$.

LEMMA 10.7. *Let $L \in \mathbb{R} \setminus \{0, -1, -2, \dots\}$. Then uniformly over $0 \le x < 1$ it holds that*

$$\tilde{f}_{n,L}(x) = \frac{1}{1-x}\left[1 + O\left(\frac{1}{n(1-x)}\right)\right] \tag{10.7}$$

the implicit constant depends only on L. Furthermore, if $L \ge 1$ then uniformly over $x \in [-1, 0]$ it holds that $\tilde{f}_{n,L}(x) = O(1)$.

PROOF. For every x we have

$$\begin{aligned}
\tilde{f}_{n,L}(x) &= \sum_{k=0}^{n} \frac{L\cdots(L+n-k-1)n!}{L\cdots(L+n-1)(n-k)!}x^k \\
&= \sum_{k=0}^{n} \frac{(n-k+1)\cdots n}{(L+n-k)\cdots(L+n-1)}x^k \\
&= \sum_{k=0}^{n} \frac{x^k}{(1+\frac{L-1}{n-k+1})\cdots(1+\frac{L-1}{n})}.
\end{aligned} \tag{10.8}$$

Now it is clear that if $x \in [-1, 0]$ and $L \ge 1$ then (10.8) is an alternating sum where the terms have decreasing modulus, thus is clearly bounded above by $O(1)$.

Now we consider $L \in \mathbb{R} \setminus \{0, -1, \dots,\}$ and $x \in [0, 1)$. Notice that for $0 \le k \le n/2$ (and n large) it holds that $0 < 1 - \frac{2|L-1|}{n} \le 1 + \frac{L-1}{n-k+1} \le 1 + \frac{2|L-1|}{n}$. It follows that $(1+\frac{L-1}{n-k+1})\cdots(1+\frac{L-1}{n}) \approx 1$, therefore, by a telescoping argument we obtain

$$\frac{1}{(1+\frac{L-1}{n-k+1})\cdots(1+\frac{L-1}{n})} = 1 + O\left(\frac{k}{n}\right)$$

(the implicit constant depends on L). Consequently, the sum of the first $n/2$ terms of $\tilde{f}_{n,L}$ satisfies

$$\begin{aligned}
\sum_{0 \le k \le n/2} \frac{x^k}{(1+\frac{L-1}{n-k+1})\cdots(1+\frac{L-1}{n})} &= \sum_{0 \le k \le n/2} x^k + \frac{1}{n}O\left(\sum_{k \ge 0} kx^k\right) \\
&= \frac{1}{1-x} + O\left(\frac{1}{n(1-x)^2}\right).
\end{aligned}$$

For the other terms, we use the classical estimate

$$C_0 k^{L-1} \le \left|\frac{L(L+1)\cdots(L+k-1)}{k!}\right| \le C_2 k^{L-1}$$

for some $C_0, C_2 > 0$ depending only on L (this estimate requires $L \notin \{0, -1, -2, \dots\}$). It follows that

$$\left| \sum_{n/2<k\le n} \frac{L(L+1)\cdots(L+n-k-1)/(n-k)!}{L(L+1)\cdots(L+n-1)/n!} x^k \right|$$
$$\le Cn^{1-L} x^{n/2} \sum_{n/2<k\le n} (n-k)^{L-1}$$
$$\le Cn^{1-L} x^{n/2} n^L \le C \frac{1}{n(1-x)^2}.$$

This completes the proof of the lemma. □

Now recall the definition of g in (10.1), we obtain

$$\tilde{f}_n(x) = \frac{1}{b_{n,L_d}} x^n g\left(\frac{1}{x}\right) + \tilde{f}_{n,L_d}(x) \tag{10.9}$$

and we have the crude estimate [which holds uniformly over $x = O(1)$)

$$\left| \frac{1}{b_{n,L_d}} x^n g(1/x) \right| \le Cn^{1-L_d}(x^n + x^{n-N_0}) \le C \frac{1}{n^{L_d+1}(1-x)^2}$$

therefore, using Lemma 10.7, we obtain the following corollary.

COROLLARY 10.8. *Uniformly, over* $0 \le x < 1$, *it holds that*

$$\tilde{f}_n(x) = \frac{1}{1-x}\left[1 + O\left(\frac{1}{n(1-x)}\right)\right].$$

Thus, using $L_d > 0$, for every fixed $x \in [0, 1)$ we have $\lim_{n\to\infty} \tilde{f}_n(x) = \frac{1}{1-x} \equiv \tilde{f}_\infty(x)$ as claimed earlier. Furthermore, from Corollary 10.8, it follows that for any fixed $c > 0$, if $x \in [0, 1 - \frac{c}{n}]$ then

$$\tilde{u}_n(x) := \frac{\tilde{f}_n(x)}{\tilde{f}_\infty(x)} = 1 + O\left(\frac{1}{c}\right)$$

for therefore by choosing $c > 0$ sufficiently large we could ensure that $\tilde{u}_n(x) \ge 1/2$ for every $x \in [0, 1 - \frac{c}{n}]$ and every n sufficiently large. Thus, by Theorem 9.2 and Lemma 9.3, we obtain the following estimate, uniformly over $t \in [\frac{1}{2}, 1 - \frac{c}{n}]$:

$$\tilde{\rho}_n(t) = \frac{1}{2\pi(1-t)} + O(1) + O\big(|\tilde{u}_n'(t^2)|^{1/2} + |\tilde{u}_n'(t^2)| + |\tilde{u}_n''(t^2)|^{1/2}\big).$$

Thus, to complete the proof of Lemma 10.6, it remains to show the following estimates uniformly over $x \in [\frac{1}{4}, 1 - \frac{c}{n}]$:

$$\tilde{u}_n'(x) = O\left(\frac{1}{n(1-x)^2}\right), \tag{10.10}$$

$$\tilde{u}_n''(x) = O\left(\frac{1}{n(1-x)^3}\right). \tag{10.11}$$

Recall from (10.9) that

$$\tilde{f}_n(x) = \frac{1}{b_{n,L_d}} x^n g\left(\frac{1}{x}\right) + \tilde{f}_{n,L_d}(x).$$

Using the definition (10.1) for g and using $L_d \geq 0$, it follows that the first term $g_1(x) := \frac{1}{b_{n,L_d}} x^n g(\frac{1}{x})$ satisfies

$$\frac{d}{dx} g_1(x) = O(n^{2-L_d} x^n) = O\left(\frac{1}{n^{L_d+1}(1-x)^3}\right) = O\left(\frac{1}{n(1-x)^2} \tilde{f}_\infty(x)\right),$$
$$\left(\frac{d}{dx}\right)^2 g_1(x) = O(n^{3-L_d} x^n) = O\left(\frac{1}{n^{L_d+1}(1-x)^4}\right) = O\left(\frac{1}{n(1-x)^3} \tilde{f}_\infty(x)\right),$$

uniformly over $x \in [1/2, 1)$.

Therefore, it suffices to show (10.10) and (10.11) for $f_n = f_{n,L_d}$. We will use the following analogue of Lemma 10.5.

LEMMA 10.9. *For $L \notin \{0, -1, -2, \dots\}$, it holds that*

$$\tilde{f}_{n,L}(x) = \frac{1}{1-x} - \frac{x}{1-x} \frac{L-1}{L+n-1} \tilde{f}_{n,L-1}(x).$$

PROOF. This follows from Lemma 10.5 using the definition of $\tilde{f}$. Alternatively, we could directly compute

$$\begin{aligned}
(1-x)\tilde{f}_{n,L}(x) &= 1 + \sum_{k=1}^{n} \frac{b_{n-k,L} - b_{n-k+1,L}}{b_{n,L}} x^k - \frac{1}{b_{n,L}} x^{n+1} \\
&= 1 - \sum_{k=1}^{n} \frac{b_{n-k+1,L-1}}{b_{n,L}} x^k - \frac{1}{b_{n,L}} x^{n+1} \\
&= 1 - \frac{L-1}{L+n-1} \sum_{k=1}^{n} \frac{b_{n-k+1,L-1}}{b_{n,L-1}} x^k - \frac{L-1}{L+n-1} \frac{1}{b_{n,L-1}} x^{n+1} \\
&= 1 - \frac{L-1}{L+n-1} x \tilde{f}_{n,L-1}(x)
\end{aligned}$$

giving the desired claim. □

Using Lemma 10.7 and Lemma 10.9, it follows that if $L \notin \{1, 0, -1, \dots\}$ then

$$\tilde{f}_{n,L}(x) = \frac{1}{1-x} - \frac{L-1}{L+n-1} \frac{x}{(1-x)^2} + O\left(\frac{1}{n^2(1-x)^3}\right). \tag{10.12}$$

On the other hand, if $L = 1$ then this estimate holds trivially via explicit computation from $\tilde{f}_{n,1} = (1 - x^{n+1})/(1 - x)$. Thus, (10.12) holds for any $L \notin \{0, -1, -2, \ldots\}$.

Now, using (10.12) and Lemma 10.9 again, we obtain the following corollary.

COROLLARY 10.10. *For $x \in [1/2, 1 - c/n]$, if $L \notin \{0, -1, -2, \ldots\}$, then*

$$\begin{aligned}\tilde{f}_{n,L}(x) &= \frac{1}{1-x} - \frac{L-1}{L+n-1}\frac{x}{(1-x)^2} + O\left(\frac{1}{n^2(1-x)^3}\right)\\ &= \frac{1}{1-x} - \frac{(L-1)}{(L+n-1)}\frac{x}{(1-x)^2}\\ &\quad + \frac{(L-1)(L-2)}{(L+n-1)(L+n-2)}\frac{x^2}{(1-x)^3} + O\left(\frac{1}{n^3(1-x)^4}\right).\end{aligned}$$

[Again, the case $L = 1$ of the second estimate in Corollary 10.10 does not follow from (10.12) and Lemma 10.9 and one checks this case separately using explicit computation.]

Now we show the desired estimate (10.10) for $\tilde{u}_n'$. As remarked earlier, it suffices to assume $f_n = f_{n,L}$ for some $L > 1$. We have

$$\tilde{u}_n'(x) = \big(\tilde{f}_{n,L}(x)(1-x)\big)' = (1-x)\frac{d}{dx}\tilde{f}_{n,L}(x) - \tilde{f}_{n,L}(x),$$

$$\frac{d}{dx}\tilde{f}_{n,L}(x) = \frac{1}{b_{n,L}}\left[nx^{n-1}f_{n,L}\left(\frac{1}{x}\right) - x^{n-2}f_{n,L}'\left(\frac{1}{x}\right)\right].$$

It is clear that $f_{n,L}'(x) = L\sum_{k=0}^{n-1}\frac{(L+1)\cdots(L+k)}{k!}x^k = Lf_{n-1,L+1}(x)$, therefore,

$$\frac{d}{dx}\tilde{f}_{n,L}(x) = \frac{n}{x}\big[\tilde{f}_{n,L}(x) - \tilde{f}_{n-1,L+1}(x)\big]. \tag{10.13}$$

Recall that $L_d > 0$. Thus, by Corollary 10.10, we have

$$\begin{aligned}\tilde{u}_n'(x) &= \frac{n(1-x)}{x}\big[\tilde{f}_{n,L_d}(x) - \tilde{f}_{n-1,L_d+1}(x)\big] - \tilde{f}_{n,L_d}(x)\\ &= \frac{n(1-x)}{x}\left[\left(\frac{1}{1-x} - \frac{L_d-1}{L_d+n-1}\frac{x}{(1-x)^2}\right)\right.\\ &\quad\left. + (-1)\left(\frac{1}{1-x} - \frac{L_d}{L_d+n-1}\frac{x}{(1-x)^2}\right)\right]\\ &\quad + \left(-\frac{1}{1-x}\right) + O\left(\frac{1}{n(1-x)^2}\right)\\ &= O\left(\frac{1}{n(1-x)^2}\right)\end{aligned}$$

uniform over $x \in [1/2, 1 - c/n]$, thus proving (10.10).

For (10.11), we use (10.13) to obtain

$$\tilde{u}_n''(x) = \frac{d}{dx}\biggl[\frac{n(1-x)}{x}\bigl(\tilde{f}_{n,L}(x) - \tilde{f}_{n-1,L+1}(x)\bigr)\biggr] - \frac{d}{dx}\tilde{f}_{n,L}(x)$$
$$= \frac{n(1-x)}{x}\biggl[\frac{n}{x}\bigl(\tilde{f}_{n,L}(x) - \tilde{f}_{n-1,L+1}(x)\bigr) - \frac{n-1}{x}\bigl(\tilde{f}_{n-1,L+1}(x) + (-1)\tilde{f}_{n-2,L+2}(x)\bigr)\biggr] + \biggl(-\frac{n}{x^2} - \frac{n}{x}\biggr)\bigl[\tilde{f}_{n,L}(x) - \tilde{f}_{n-1,L+1}(x)\bigr].$$

Using Corollary 10.10 again, we have

$$\tilde{f}_{n,L}(x) - \tilde{f}_{n-1,L+1}(x) = \frac{x}{(L+n-1)(1-x)^2} + \frac{-2(L-1)}{(L+n-1)(L+n-2)}\frac{x^2}{(1-x)^3} + O\biggl(\frac{1}{n^3(1-x)^4}\biggr).$$

Therefore,

$$\begin{aligned}
\tilde{u}_n''(x) &= \frac{n(1-x)}{x}\biggl[\frac{1}{(L_d+n-1)(1-x)^2}\biggr] \\
&\quad + \frac{n(1-x)}{x}\biggl[\frac{2L_d(n-1) - 2(L_d-1)n}{(L_d+n-1)(L_d+n-2)}\frac{x}{(1-x)^3}\biggr] \\
&\quad + (-1)\frac{n+nx}{x(L_d+n-1)(1-x)^2} + O\biggl(\frac{1}{n(1-x)^3}\biggr) \\
&= \frac{n(1-x)}{x}\biggl[\frac{1}{n(1-x)^2} + O\biggl(\frac{1}{n^2(1-x)^2}\biggr)\biggr] \\
&\quad + \frac{n(1-x)}{x}\biggl[\frac{2L_dn - 2(L_d-1)n}{n^2}\frac{x}{(1-x)^3} + O\biggl(\frac{1}{n^2(1-x)^3}\biggr)\biggr] \\
&\quad + (-1)\frac{n+nx}{nx(1-x)^2} + O\biggl(\frac{1}{n(1-x)^2}\biggr) + O\biggl(\frac{1}{n(1-x)^3}\biggr) \\
&= \frac{1}{x(1-x)} + \frac{2n}{n(1-x)^2} - \frac{n+nx}{x(1-x)^2} + O\biggl(\frac{1}{n(1-x)^3}\biggr) \\
&= O\biggl(\frac{1}{n(1-x)^3}\biggr)
\end{aligned}$$

thus proving (10.11). This completes the proof of Lemma 10.2.

11. Proof of Theorem 2.9. In this section, we count the average number of real zeros for $P_n(t) = \sum_{j=0}^{n} c_j\xi_j t^j$ where for $j \geq N_0$ two conditions hold: $c_j =$

$\mathfrak{P}(j)$ for some fixed classical polynomial $\mathfrak{P}$ of degree ρ when $j \geq N_0$ and is bounded when $j \leq N_0$, and ξ_j are independent Gaussian with mean $\mu \neq 0$ and variance 1. Without loss of generality, we may assume that the leading coefficient of $\mathfrak{P}$ is 1, that is, $c_j = j^\rho + \cdots$.

Thanks to Lemma 2.5, the average number of real zeros outside $[-1-b_1, -1+b_1]$ and $[1-b_1, 1+b_1]$ (for any fixed $b_1 > 0$) is bounded.

We will show that on average there are a bounded number of real zeros in $[1-b_1, 1+b_1]$ and $\frac{1+\sqrt{2\rho+1}}{2\pi}\log n + O(1)$ real zeros in $[-1-b_1, -1+b_1]$.

As in the proof of Corollary 2.6, let $m(t) = \mathbf{E}P(t)$, $\mathcal{P} = \mathbf{Var}[P_n] = \sum_{j=0}^n c_j^2 t^{2j}$, $\mathcal{Q} = \mathbf{Var}[P_n'(t)] = \sum_{j=0}^n c_j j^2 t^{2j-2}$, and $\mathcal{R} = \mathbf{Cov}[P_n, P_n'] = \sum_{j=0}^n j c_j^2 t^{2j-1}$, and $\mathcal{S} = \mathcal{P}\mathcal{Q} - \mathcal{R}^2$.

We will use the following generalization of the Kac–Rice formula in [12], Corollary 2.1, which gives

$$\begin{aligned}
\mathbf{E}N_n[a,b] &= I_1(a,b) + I_2(a,b),\\
I_1(a,b) &:= \int_a^b \frac{\mathcal{S}^{1/2}}{\pi\mathcal{P}} \exp\left(-\frac{m^2\mathcal{Q} + m'^2\mathcal{P} - 2mm'\mathcal{R}}{2\mathcal{S}}\right) dt,\\
I_2(a,b) &:= \int_a^b \frac{\sqrt{2}|m'\mathcal{P} - m\mathcal{R}|}{\pi\mathcal{P}^{3/2}} \exp\left(-\frac{m^2}{2\mathcal{P}}\right) erf\left(\frac{|m'\mathcal{P} - m\mathcal{R}|}{\sqrt{2\mathcal{P}\mathcal{S}}}\right) dt,\\
erf(x) &:= \int_0^x e^{-t^2}\, dt.
\end{aligned}$$

We note that in I_1 the first factor $\mathcal{S}^{1/2}/(\pi\mathcal{P})$ is exactly the density of the real roots for P_n in the mean zero case, namely ρ_n in the notation of Lemma 9.1, and there is an extra exponential factor in I_1. Our plan is, essentially, to show that near 1 the exponential decay of the extra factor in I_1 will cancel out the pole singularity of ρ_n and near -1 the extra factor in I_1 is essentially 1. This would lead to $I_1(a,b) = O(1)$ if a, b are close to 1 and $I_1(a,b) = \int_a^b \rho_n(t)\, dt + O(1)$ if a, b, are close to -1, thus allowing us to reduce the proof to the mean zero case. For I_2, we will show that $I_2(a,b) = O(1)$ for both cases.

We now separate the neighborhood into four intervals: $[1-b_1, 1]$, $[-1, -1+b_1]$, $[1, 1+b_1]$, and $[-1-b_1, -1]$, where $b_1 > 0$ is a sufficiently small fixed constant.

The interval $[1-b_1, 1]$. We will show that this interval contributes $O(1)$ to $\mathbf{E}N_n$. Using (10.3), for $1 - b_1 \leq t < 1$ we have

$$\begin{aligned}
m(t) &= \mu \sum_{j=0}^n c_j t^j\\
&= \sum_{j<N_0} \mu[c_j - \mathfrak{P}(j)]t^j + \frac{\mu\rho!}{(1-t)^{\rho+1}}\left(1 + O\left([1+n(1-t)]^\rho t^{n+1}\right)\right)
\end{aligned}$$

$$= O(1) + \frac{\mu\rho!}{(1-t)^{\rho+1}}\bigl(1 + O\bigl([1+n(1-t)]^{\rho}t^{n+1}\bigr)\bigr)$$
$$= O(1) + \frac{\mu\rho!}{(1-t)^{\rho+1}}\bigl(1 + O\bigl(t^{n/2}\bigr)\bigr)$$

here we have used the fact that $n^L s^n = O((1-s)^{-L})$, applied to $s = t^{1/2}$. Similarly,

$$m'(t) = \mu\sum_{j=0}^{n-1} c_{j+1}(j+1)t^j = O(1) + \frac{\mu(\rho+1)!}{(1-t)^{\rho+2}}\bigl(1 + O\bigl(t^{n/2}\bigr)\bigr),$$
$$\mathcal{P} = O(1) + \frac{(2\rho)!}{(1-t^2)^{2\rho+1}}\bigl(1 + O\bigl(t^n\bigr)\bigr),$$
$$\mathcal{Q} = O(1) + \frac{(2\rho+2)!}{(1-t^2)^{2\rho+3}}\bigl(1 + O\bigl(t^n\bigr)\bigr),$$
$$\mathcal{R} = O(1) + \frac{(2\rho+1)!t}{(1-t^2)^{2\rho+2}}\bigl(1 + O\bigl(t^n\bigr)\bigr).$$

Note that by choosing $c > 0$ sufficiently large we could ensure that $t^{n/2} \ll 1$ for $|t| \le 1 - \frac{c}{n}$, and by choosing $b_1 > 0$ sufficiently small we could ensure that $\frac{1}{1-t^2} \gg 1$ for $t \in [1-b_1, 1)$. It follows that

$$m^2\mathcal{Q} + m'^2\mathcal{P} - 2mm'\mathcal{R} \ge C_\rho \frac{1}{(1-t)^{4\rho+5}} \ge C'_\rho \frac{\mathcal{P}^2}{(1-t)^3}$$

for some positive constants C'_ρ, C_ρ depending only on ρ and μ. Now, by Lemma 10.3, we have

$$\frac{\mathcal{S}^{1/2}}{\pi\mathcal{P}} = \rho_n \sim \frac{1}{1-|t|}.$$

Consequently, uniformly over $t \in [1-b_1, 1]$, we have

$$\frac{m^2\mathcal{Q} + m'^2\mathcal{P} - 2mm'\mathcal{R}}{2\mathcal{S}} \ge C''_\rho \frac{1}{1-t}$$

therefore,

$$I_1\left(1-b_1, 1-\frac{c}{n}\right) = O\left(\int_{1-b_1}^{1} \frac{1}{1-t}\exp\left(-\frac{C''_\rho}{1-t}\right)dt\right) = O(1)$$
$$I_1\left(1-\frac{c}{n}, 1\right) \le \int_{1-\frac{c}{n}}^{1} \rho_n(t)\,dt = O(1).$$

Now for I_2 we similarly have, for $t \in [1-b_1, 1-\frac{c}{n}]$,

$$\frac{m^2}{2\mathcal{P}} \ge C_\rho \frac{1}{1-t}, \qquad \bigl|m'\mathcal{P} - m\mathcal{R}\bigr| = O\left(\frac{1}{(1-t)^{3\rho+3}}\right) = O\left(\frac{\mathcal{P}^{3/2}}{(1-t)^{3/2}}\right),$$

therefore,

$$I_2\left(1-b_1, 1-\frac{c}{n}\right) = O\left(\int_{1-b_1}^{1-\frac{c}{n}} \frac{1}{(1-t)^{3/2}} \exp\left(-C_\rho \frac{1}{1-t}\right) dt\right) = O(1).$$

On the other hand, the integrand of I_2 is bounded above by $O(n)$ for $t \in [1-\frac{c}{n}, 1]$, for any fixed $c > 0$. To see this, first note that for some absolute constant n_0 the coefficients c_j are of the same sign and $|c_j| \geq j^\rho$ for $j \geq n_0$. It follows that the main contribution to m and m' comes from the tail $j \geq n_0$. For instance,

$$|m(t)| = O(1) + \left|\sum_{n_0 \leq j \leq n} c_j t^j\right|,$$

$$\left|\sum_{n_0 \leq j \leq n} c_j t^j\right| \geq \frac{1}{C} \sum_{n_0 \leq j \leq n} j^\rho \geq \frac{1}{C'} n^{\rho+1} \gg 1,$$

and for m' we could argue similarly. Since c_j are of the same sign for $j \geq n_0$, it follows immediately that $|m'(t)| = O(n|m(t)|)$, and consequently

$$\frac{|m'\mathcal{P}|}{\mathcal{P}^{3/2}} \exp\left(-\frac{m^2}{2\mathcal{P}}\right) = nO\left(\frac{|m|}{\mathcal{P}^{1/2}} \exp\left(-\frac{m^2}{2\mathcal{P}}\right)\right) = O(n),$$

using the boundedness of xe^{-x^2}. We also have

$$\frac{|m\mathcal{R}|}{\mathcal{P}^{3/2}} \exp\left(-\frac{m^2}{2\mathcal{P}}\right) = O\left(\frac{\mathcal{R}}{\mathcal{P}}\right) = O(n).$$

It follows that $I_2(1-\frac{c}{n}, 1) = O(1)$, so $I_2(1-b_1, 1) = O(1)$.

The interval $[-1, -1+b_1]$. We will show that this interval contributes $(\sqrt{2\rho+1}\log n)/(2\pi) + O(1)$ to $\mathbf{E}N_n$. The analysis of this interval is fairly similar to the analysis of $[1-b_1, 1]$; the main difference is that $m(t)$ and $m'(t)$ are less singular near -1, in fact they are bounded by $O((1+t)^{-\rho})$ and $O((1+t)^{-(\rho+1)})$, respectively [by using (10.4) for $L = 1, 2, \dots, \rho$ and expanding the polynomial defining c_j into the linear basis of binomial polynomials]. It follows that

$$m^2\mathcal{Q} + m'^2\mathcal{P} - 2mm'\mathcal{R} = O\left(\frac{1}{(1+t)^{4\rho+3}}\right) = O\left(\frac{\mathcal{P}^2}{1+t}\right) = O\big((1+t)\mathcal{S}\big)$$

therefore, for $c > 0$ sufficiently large and $b_1 > 0$ sufficiently small

$$\begin{aligned} I_1\left(-1+\frac{c}{n}, -1+b_1\right) &= \int_{-1+\frac{c}{n}}^{-1+b_1} \rho_n(t)\,dt + O\left(\int_{-1}^{-1+b_1} \rho_n(t)(1+t)\,dt\right) \\ &= \frac{\sqrt{2\rho+1}}{2\pi}\log n + O(1), \\ I_1\left(-1, -1+\frac{c}{n}\right) &\leq \int_{-1}^{-1+\frac{c}{n}} \rho_n(t)\,dt = O(1). \end{aligned}$$

For I_2, similarly we only need to show that $I_2(-1+\frac{c}{n}, -1+b_1) = O(1)$. This follows from

$$\frac{|m'\mathcal{P} - m\mathcal{R}|}{\mathcal{P}^{3/2}} = O\left(\frac{(1+t)^{-(2\rho+2)}}{(1+t)^{-3\rho+\frac{3}{2}}}\right) = O\big((1+t)^{\rho-\frac{1}{2}}\big) = O\big((1+t)^{-\frac{1}{2}}\big).$$

The interval $[1, 1+b_1]$. We will show that this interval contributes $O(1)$ to $\mathbf{E}N_n$. To analyze this interval, consider the reciprocal polynomial $\widetilde{P}_n(t) = \sum_{j=0}^n \tilde{c}_j \xi_j t^j$ where $\tilde{c}_j = c_{n-j}/c_n$. For convenience, let $\tilde{\rho}_n$, $\widetilde{I}_1$, $\widetilde{I}_2$, $\widetilde{\mathcal{P}}$, $\widetilde{\mathcal{Q}}$, $\widetilde{\mathcal{R}}$, $\widetilde{\mathcal{S}}$, $\tilde{m}$, and $\tilde{m}'$ be the corresponding quantities, and similarly it suffices to show that $\widetilde{I}_1(1-b_1, 1-\frac{c}{n})$, $\widetilde{I}_2(1-b_1, 1-\frac{c}{n}) = O(1)$ where $c > 0$ is a fixed large constant.

Let $\tilde{f}_n(t) = \widetilde{\mathcal{P}}(t)$, so $\tilde{f}_n(t^2) = \mathbf{Var}[\widetilde{P}_n(t)]$ as in the proof of Theorem 2.8. Recall from the proof of Lemma 9.1 that $\widetilde{\mathcal{Q}} = \sum_{j=0}^n j^2 \tilde{c}_j^2 t^{2j-2} = \tilde{f}_n'(t^2) + t^2 \tilde{f}_n''(t^2)$, and $\widetilde{\mathcal{R}} = \sum_{j=0}^n j \tilde{c}_j^2 t^{2j-1} = t \tilde{f}_n'(t^2)$.

Recall that $\tilde{u}_n(x) = \tilde{f}_n(x)(1-x)$. From Corollary 10.8, (10.10) and (10.11), for $x \in [1-b_1, 1-\frac{c}{n}]$ with $c > 0$ sufficiently large we have

$$\begin{aligned}
\tilde{f}_n(x) &= \frac{1}{1-x}\left[1 + O\left(\frac{1}{n(1-x)}\right)\right], \\
\tilde{f}_n'(x) &= \frac{\tilde{u}_n'(x) + \tilde{f}_n(x)}{1-x} \\
&= \frac{\tilde{f}_n(x)}{1-x} + O\left(\frac{1}{n(1-x)^3}\right) \\
&= \frac{1}{(1-x)^2} + O\left(\frac{1}{n(1-x)^3}\right), \\
\tilde{f}_n''(x) &= \frac{\tilde{u}_n''(x) + 2\tilde{f}_n'(x)}{1-x} = \frac{2}{(1-x)^3} + O\left(\frac{1}{n(1-x)^4}\right).
\end{aligned}$$

It follows that for $t \in [1-b_1, 1-\frac{c}{n}]$, we have

$$\begin{aligned}
\widetilde{\mathcal{P}}(t) &= \frac{1}{1-t^2} + O\left(\frac{1}{n(1-t^2)^2}\right), \\
\widetilde{\mathcal{Q}}(t) &= \frac{2}{(1-t^2)^3} + O\left(\frac{1}{n(1-t^2)^4}\right) + O\left(\frac{1}{(1-t^2)^2}\right), \\
\widetilde{\mathcal{R}}(t) &= \frac{t}{(1-t^2)^2} + O\left(\frac{1}{n(1-t^2)^3}\right)
\end{aligned}$$

and using Lemma 10.6 and the Edelman–Kostlan formula we have

$$\widetilde{\mathcal{S}} = \tilde{\rho}_n(t)^2 \pi^2 \widetilde{\mathcal{P}}^2(t) = O\left(\frac{1}{(1-t^2)^2}\widetilde{\mathcal{P}}^2\right).$$

On the other hand, for $t \in [1-b_1, 1]$, using Corollary 10.8 we have

$$\tilde{m}(t) = \mu \sum_{j=0}^{n} \tilde{c}_j t^j = \frac{\mu}{1-t}\left[1 + O\left(\frac{1}{n(1-t)}\right)\right].$$

Let $d_j = c_j j$ which is a polynomial of j (for $j \geq N_0$) of degree $\rho + 1$. Then for $j \leq n - N_0$ we have $\tilde{d}_j = \frac{d_{n-j}}{d_n} = \tilde{c}_j - \frac{j}{n}\tilde{c}_j$. We obtain

$$\tilde{m}'(t) = \mu \sum_{j=0}^{n} \tilde{c}_j j t^{j-1} = n\mu \sum_{j=0}^{n} \tilde{c}_j t^{j-1} - n\mu \sum_{j=0}^{n} \tilde{d}_j t^{j-1}.$$

To evaluate $\sum_{j=0}^{n} \tilde{d}_j t^{j-1}$ and $\sum_{j=0}^{n} \tilde{c}_j t^{j-1}$, we use Corollary 10.10 together with an expansion of the polynomials defining c_j and d_j into the linear basis of binomial polynomials $\frac{L\cdots(L+j-1)}{j!}$ with $L = 1, 2, \dots$ (as in the proof of Corollary 10.8). It follows that

$$\begin{aligned}
\tilde{m}'(t) &= \mu \sum_{j=0}^{n} \tilde{c}_j j t^{j-1} \\
&= n\mu \sum_{j=0}^{n} \tilde{c}_j t^{j-1} - n\mu \sum_{j=0}^{n} \tilde{d}_j t^{j-1} \\
&= n\mu\left[\frac{1}{1-t}\left(1 + O\left(\frac{1}{n}\right)\right) - \frac{\rho}{\rho+n}\frac{t}{(1-t)^2} + O\left(\frac{1}{n^2(1-t)^2}\right)\right] \\
&\quad - n\mu\left[\frac{1}{1-t}\left(1 + O\left(\frac{1}{n}\right)\right) - \frac{\rho+1}{\rho+1+n}\frac{t}{(1-t)^2} + O\left(\frac{1}{n^2(1-t)^2}\right)\right] \\
&= \frac{\mu}{(1-t)^2} + O\left(\frac{1}{1-t}\right).
\end{aligned}$$

Note that by choosing b_1 small and c large we know that $1-t \ll 1$ and $\frac{1}{n(1-t)} \ll 1$. Thus,

$$\begin{aligned}
\tilde{m}^2\widetilde{\mathcal{Q}} + \tilde{m}'^2\widetilde{\mathcal{P}} - 2\tilde{m}\tilde{m}'\widetilde{\mathcal{R}} &\geq C^{-1}\frac{1}{(1-t)^5} \geq \widetilde{\mathcal{P}}^2(1-t)^{-3} \\
&\geq C^{-1}\widetilde{\mathcal{S}}(1-t)^{-1}
\end{aligned}$$

and the rest of the proof is similar to the prior treatment for (the case $\rho = 0$ of) $\mathbf{E}N_n[1-b_1, 1]$. In particular, to show that $\tilde{m}'(t) = O(n|\tilde{m}(t)|)$ for $t \in [1 - \frac{c}{n}, 1]$ [in the treatment of $\widetilde{I}_2(1-\frac{c}{n}, 1)$] we similarly observe that the main contributions to $|\tilde{m}|$ and $|\tilde{m}'|$ come from $0 \leq j \leq n - n_0$, and for these indices we have $\tilde{c}_j > 0$.

The interval $[-1-b_1, -1]$. We will show that this interval contributes $(\log n)/(2\pi) + O(1)$ to $\mathbf{E}N_n$. As before, we also consider the reciprocal polynomial $\widetilde{P}_n$ and count the number of real roots in $[-1, -1+b_1]$ for this polynomial. The analysis is similar to the treatment for the interval $[1, 1+b_1]$; the only modification is in the estimate for $\tilde{m}$ and $\tilde{m}'$ near -1, and unlike the last interval here these two terms are bounded above by $O(1)$ (via applications of Lemma 10.7 together with an expansion of the polynomial defining c_j into the linear basis of binomial polynomials). The rest of the proof is entirely similar to the prior treatment for (the case $\rho = 0$ of) $\mathbf{E}N_n[-1, -1 + b_1]$.

APPENDIX

In this section, we provide the proof of Theorem 4.11. For the proof of Theorem 2.10, we need an analog of this theorem when P is a power series of the form (2.11). A proof for series in fact runs along the same line with the following proof for polynomials, except some minor modifications that we shall notify the reader.

We first prove the following lemma.

LEMMA A.1. *Let P be the random polynomial of the form (2.4) where the ξ_i are independent random variables with variance 1 and $\sup_{i\geq 0} \mathbf{E}|\xi_i|^{2+\epsilon} \leq \tau_2$ for some constant τ_2. And let $\widetilde{P} = \sum_{i=0}^{\infty} c_i \tilde{\xi}_i z^i$ be the corresponding polynomial with Gaussian random variables $\tilde{\xi}_i$. Assume that $\tilde{\xi}_i$ matches moments to second order with ξ_i for every $i \in \{0, \dots, n\} \setminus I_0$ for some subset I_0 (may depend on n) of size bounded by some constant N_0 and that $\sup_{i\geq 0} \mathbf{E}|\tilde{\xi}_i|^{2+\epsilon} \leq \tau_2$.*

Then there exists a constant C_2 such that the following holds true. Let $\alpha_1 \geq C_2\alpha_0 > 0$ and $C > 0$ be any constants. Let $\delta \in (0, 1)$ and $m \leq \delta^{-\alpha_0}$ and $z_1, \dots, z_m \in \mathbb{C}$ be complex numbers such that

$$\frac{|c_i||z_j|^i}{\sqrt{V(z_j)}} \leq C\delta^{\alpha_1} \qquad \forall i = 0, \dots, n, j = 1, \dots, m, \tag{A.1}$$

where $V(z_j) = \sum_{i=\{0,\dots,n\}\setminus I_0} |c_i|^2 |z_j|^{2i}$. Let $H : \mathbb{C}^m \to \mathbb{C}$ be any smooth function such that $\|\nabla^a H\| \leq \delta^{-\alpha_0}$ for all $0 \leq a \leq 3$, then

$$\left|\mathbf{E}H\left(\frac{P(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P(z_m)}{\sqrt{V(z_m)}}\right) - \mathbf{E}H\left(\frac{\widetilde{P}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{\widetilde{P}(z_m)}{\sqrt{V(z_m)}}\right)\right| \leq \widetilde{C}\delta^{\alpha_0},$$

where $\widetilde{C}$ is a constant depending only on $\alpha_0, \alpha_1, C, N_0, \tau_2$ and not on δ.

PROOF. Our proof works for any subset I_0 of size bounded by N_0, but for notation convenience, we assume that $I_0 = \{0, \dots, N_0 - 1\}$. We use the Lindeberg swapping argument. Let $P_{i_0} = \sum_{i=0}^{i_0-1} c_i \tilde{\xi}_i z^i + \sum_{i=i_0}^{n} c_i \xi_i z^i$. Then $P_0 = P$ and $P_{n+1} = \widetilde{P}$. Put

$$I_{i_0} = \left|\mathbf{E}H\left(\frac{P_{i_0}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P_{i_0}(z_m)}{\sqrt{V(z_m)}}\right) - \mathbf{E}H\left(\frac{P_{i_0+1}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P_{i_0+1}(z_m)}{\sqrt{V(z_m)}}\right)\right|.$$

Then[6]

$$I := \left|\mathbf{E}H\left(\frac{P(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P(z_m)}{\sqrt{V(z_m)}}\right) - \mathbf{E}H\left(\frac{\widetilde{P}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{\widetilde{P}(z_m)}{\sqrt{V(z_m)}}\right)\right| \leq \sum_{i_0=0}^{n} I_{i_0}.$$

Fix $i_0 \geq N_0$ and let $Y_j = \sum_{i=0}^{i_0-1} \frac{c_i \tilde{\xi}_i z_j^i}{\sqrt{V(z_j)}} + \sum_{i=i_0+1}^{n} \frac{c_i \xi_i z_j^i}{\sqrt{V(z_j)}}$ for j from 1 to n. Then $\frac{P_{i_0}(z_j)}{\sqrt{V(z_j)}} = Y_j + \frac{c_{i_0}\xi_{i_0} z_j^{i_0}}{\sqrt{V(z_j)}}$ and $\frac{P_{i_0+1}(z_j)}{\sqrt{V(z_j)}} = Y_j + \frac{c_{i_0}\tilde{\xi}_{i_0} z_j^{i_0}}{\sqrt{V(z_j)}}$. Fix ξ_i when $i < i_0$ and $\tilde{\xi}_i$ when $i > i_0$ and the Y_j's are fixed. Put

$$G = G_{i_0}(w_1, \dots, w_m) := H(Y_1 + w_1, \dots, Y_m + w_m).$$

Then $\|\nabla^a G\|_\infty \leq C\delta^{-\alpha_0}$ for all $0 \leq a \leq 3$. Then we need to estimate d_{i_0}, which is defined by the following expression:

$$\left|\mathbf{E}_{\xi_{i_0},\tilde{\xi}_{i_0}} G\left(\frac{c_{i_0}\xi_{i_0} z_1^{i_0}}{\sqrt{V(z_1)}}, \dots, \frac{c_{i_0}\xi_{i_0} z_m^{i_0}}{\sqrt{V(z_m)}}\right) - \mathbf{E}_{\xi_{i_0},\tilde{\xi}_{i_0}} G\left(\frac{c_{i_0}\tilde{\xi}_{i_0} z_1^{i_0}}{\sqrt{V(z_1)}}, \dots, \frac{c_{i_0}\tilde{\xi}_{i_0} z_m^{i_0}}{\sqrt{V(z_m)}}\right)\right|.$$

Let $a_{i,i_0} = \frac{c_{i_0} z_i^{i_0}}{\sqrt{V(z_i)}}$ and $a_{i_0} = (\sum_{i=1}^m |a_{i,i_0}|^2)^{1/2}$. Taylor expanding G around $(0, \dots, 0)$ gives

$$G(a_{1,i_0}\xi_{i_0}, \dots, a_{m,i_0}\xi_{i_0}) = G(0) + G_1 + \mathrm{err}_1, \tag{A.2}$$

where

$$G_1 = \frac{dG(a_{1,i_0}\xi_{i_0}t, \dots, a_{m,i_0}\xi_{i_0}t)}{dt}\bigg|_{t=0} = \sum_{i=1}^m \frac{\partial G(0)}{\partial\,\mathrm{Re}(w_i)} \mathrm{Re}(a_{i,i_0}\xi_{i_0}) + \sum_{i=1}^m \frac{\partial G(0)}{\partial\,\mathrm{Im}(w_i)} \mathrm{Im}(a_{i,i_0}\xi_{i_0})$$

[6]For power series, to have $I \leq \sum_{i_0=0}^{\infty} I_{i_0}$, we need to show that

$$\mathbf{E}H\left(\frac{P_0(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P_0(z_m)}{\sqrt{V(z_m)}}\right) - \mathbf{E}H\left(\frac{\widetilde{P}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{\widetilde{P}(z_m)}{\sqrt{V(z_m)}}\right) = \sum_{i_0=0}^{\infty}\left(\mathbf{E}H\left(\frac{P_{i_0}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P_{i_0}(z_m)}{\sqrt{V(z_m)}}\right) - \mathbf{E}H\left(\frac{P_{i_0+1}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P_{i_0+1}(z_m)}{\sqrt{V(z_m)}}\right)\right),$$

that is, $\mathbf{E}H(\frac{P_n(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{P_n(z_m)}{\sqrt{V(z_m)}}) \to \mathbf{E}H(\frac{\widetilde{P}(z_1)}{\sqrt{V(z_1)}}, \dots, \frac{\widetilde{P}(z_m)}{\sqrt{V(z_m)}})$ as $n \to \infty$. This follows from the fact that $P_n(z_i) \to \widetilde{P}(z_i)$ a.e., the continuity and boundedness of H, and the dominated convergence theorem.

and

$$|\text{err}_1| \le \sup_{t' \in [0,1]} \left| \frac{1}{2} \frac{d^2 G(a_{1,i_0}\xi_{i_0} t, \ldots, a_{m,i_0}\xi_{i_0} t)}{dt^2} \right|_{t=t'}$$

$$= \sup_{t' \in [0,1]} \left| \frac{1}{2} \sum_{h,k \in \{\text{Re},\text{Im}\}, i,j \in \{1,\ldots,m\}} \frac{\partial^2 G}{\partial h(w_i) \partial k(w_j)} h(a_{i,i_0}\xi_{i_0}) k(a_{j,i_0}\xi_{i_0}) \right|$$

$$\le \widetilde{C}\delta^{-\alpha_0} |\xi_{i_0}|^2 \sum_{i,j=1}^m |a_{i,i_0}||a_{j,i_0}| \le \widetilde{C}\delta^{-\alpha_0} |\xi_{i_0}|^2 \left(\sum_{i=1}^m |a_{i,i_0}| \right)^2$$

$$\le \widetilde{C}\delta^{-\alpha_0} |\xi_{i_0}|^2 m \left(\sum_{i=1}^m |a_{i,i_0}|^2 \right) = \widetilde{C}\delta^{-2\alpha_0} |\xi_{i_0}|^2 a_{i_0}^2.$$

Similarly,

$$G(a_{1,i_0}\xi_{i_0}, \ldots, a_{m,i_0}\xi_{i_0}) = G(0) + G_1 + \frac{1}{2} G_2 + \text{err}_2, \tag{A.3}$$

where $G_2 = \frac{d^2 G(a_{1,i_0}\xi_{i_0} t, \ldots, a_{m,i_0}\xi_{i_0} t)}{dt^2}|_{t=0}$ and

$$|\text{err}_2| \le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} |\xi_{i_0}|^3 a_{i_0}^3. \tag{A.4}$$

Also, we have $|\text{err}_2| = |\text{err}_1 - \frac{1}{2}G_2| \le \widetilde{C}\delta^{-2\alpha_0}|\xi_{i_0}|^2 a_{i_0}^2 \le \delta^{-\frac{5}{2}\alpha_0}|\xi_{i_0}|^2 a_{i_0}^2$. Interpolation gives

$$|\text{err}_2| \le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} |\xi_{i_0}|^{2+\epsilon} a_{i_0}^{2+\epsilon}.$$

The expression (A.4) also holds for $\tilde{\xi}$ in place of ξ. Subtracting and taking expectations and using the matching moments give

$$d_{i_0} = |\mathbf{E}\text{err}_2| \le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} a_{i_0}^{2+\epsilon} (\mathbf{E}|\xi_{i_0}|^{2+\epsilon} + \mathbf{E}|\tilde{\xi}_{i_0}|^{2+\epsilon}) \le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} a_{i_0}^{2+\epsilon}.$$

Taking expectation with respect to the random variables $\tilde{\xi}_i$ where $i < i_0$ and ξ_i where $i > i_0$ gives $I_{i_0} \le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} a_{i_0}^{2+\epsilon}$, for all $i_0 \ge N_0$.

For $0 \le i_0 < N_0$, instead of (A.2) and (A.3), we use mean value theorem to get the rough bound

$$G(a_{1,i_0}\xi_{i_0}, \ldots, a_{m,i_0}\xi_{i_0}) = G(0) + O\left(m \|\nabla G\|_\infty |\xi_{i_0}| \sum_{i=1}^m |a_{i,i_0}| \right),$$

which by the same arguments as above gives $I_{i_0} \le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} a_{i_0}$. Thus,

$$I \le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} \sum_{i_0=0}^n a_{i_0}^{2+\epsilon} + \widetilde{C}\delta^{-\frac{5}{2}\alpha_0} \sum_{i_0=0}^{N_0} a_{i_0}.$$

Note that since $a_{i_0}^2 = \sum_{i=1}^m |c_{i_0}|^2 \frac{|z_i|^{2i_0}}{V(z_i)}$, we obtain

$$\sum_{i_0=0}^n a_{i_0}^2 = m + \sum_{i=1}^m \sum_{i_0=0}^{N_0} \frac{|c_{i_0}|^2 |z_i|^{2i_0}}{V(z_i)} = m + O(m\delta^{2\alpha_1}) = O(m).$$

Moreover, since $\frac{|c_{i_0}||z_i|^{i_0}}{\sqrt{V(z_i)}} \le \widetilde{C}\delta^{\alpha_1}$, $a_{i_0}^2 \le m\widetilde{C}^2\delta^{2\alpha_1} \le \widetilde{C}^2\delta^{2\alpha_1-\alpha_0}$. Hence,

$$\begin{aligned} I &\le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0+\epsilon(\alpha_1-\frac{\alpha_0}{2})} \sum_{i_0=0}^n a_{i_0}^2 + \widetilde{C}\delta^{\alpha_1-3\alpha_0} \\ &\le \widetilde{C}\delta^{-\frac{5}{2}\alpha_0+\epsilon(\alpha_1-\frac{\alpha_0}{2})}\delta^{-\alpha_0} + \widetilde{C}\delta^{\alpha_1-3\alpha_0} \le \widetilde{C}\delta^{\alpha_0}. \end{aligned}$$ □

Now we proceed to the proof of Theorem 4.11.

PROOF. Consider

$$\bar{F}(w_1, \dots, w_m) = F\left(w_1 + \frac{1}{2}\log|V(z_1)|, \dots, w_m + \frac{1}{2}\log|V(z_m)|\right).$$

Then we still have $\|\nabla^a \bar{F}\|_\infty \le C\delta^{-\alpha_0}$ for all $0 \le \alpha \le 3$, and we want to show that

$$\left|\mathbf{E}\bar{F}\left(\log\frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots\right) - \mathbf{E}\bar{F}\left(\log\frac{|\widetilde{P}(z_1)|}{\sqrt{V(z_1)}}, \dots\right)\right| \le \widetilde{C}\delta^{\alpha_0}$$

(the ... stop at z_m, this will be our convention when writing long expressions). Let

$$\Omega_1 = \left\{(w_1, \dots, w_m) \in \mathbb{R}^m : \min_{i=1,\dots,m} w_i < -M\right\},$$

$$\Omega_2 = \left\{(w_1, \dots, w_m) \in \mathbb{R}^m : \min_{i=1,\dots,m} w_i > -M - 1\right\},$$

where M is to be defined. Then $\Omega_1 \cup \Omega_2 = \mathbb{R}^m \subset \mathbb{C}^m$, and since we only look at

$$\bar{F}\left(\log\frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots, \log\frac{|P(z_m)|}{\sqrt{V(z_m)}}\right),$$

we can restrict $\bar{F}$ to $\mathbb{R}^m \subset \mathbb{C}^m$ and think about $\bar{F}$ as a function from $\mathbb{R}^m \to \mathbb{C}$. We can further assume that $\bar{F} : \mathbb{R}^m \to \mathbb{R}$ by considering the real and imaginary parts of $\bar{F}$ separately.

Now there exists a smooth function $\psi : \mathbb{R}^m \to \mathbb{R}$ such that ψ is supported in Ω_2 and $\psi = 1$ on the complement of Ω_1 and $\|\nabla^a \psi\|_\infty \le m^{C_2}$ for all $0 \le a \le 3$ and C_2 is some constant. Indeed, there exists a function $\rho : \mathbb{R} \to \mathbb{R}$ such that ρ is supported in $[-M-1, \infty)$, $\rho = 1$ on $[-M, \infty)$, $0 \le \rho \le 1$, and ρ has bounded derivatives of all orders. This function ρ can be constructed by convolution of the indicator of $[-M - 1/2, \infty)$ with a mollifier. Now let $\psi(x_1, \dots, x_m) = \rho(x_1)\cdots\rho(x_m)$. Then clearly ψ satisfies the required conditions.

Now put $\phi = 1 - \psi$, $F_1 = \bar{F} \cdot \phi$ and $F_2 = \bar{F}.\psi$. Then $\bar{F} = F_1 + F_2$, and both F_1, F_2 are smooth functions with $\operatorname{supp} F_1 \subset \bar{\Omega}_1$, $\operatorname{supp} F_2 \subset \bar{\Omega}_2$. We have

$$\|\nabla F_1\| = \|\nabla \bar{F}.\phi + \bar{F}\nabla\phi\| \le \|\nabla \bar{F}\|\|\phi\| + \|\bar{F}\|\|\nabla\phi\| \le \widetilde{C}\delta^{-C_2\alpha_0}.$$

And similarly for higher derivatives and for F_2, we then get $\|\nabla^a F_i\| \le \widetilde{C}\delta^{-C_2\alpha_0}$ for $i = 1, 2$ and $0 \le a \le 3$.

We now show that the contribution from F_1 is negligible. We show this by first showing that there exists a smooth function $H_1 : \mathbb{R}^m \to \mathbb{R}$ such that

$$\big|F_1\big(\log|w_1|, \dots, \log|w_m|\big)\big| \le H_1(w_1, \dots, w_m), \qquad \|\nabla^a H_1\| \le \widetilde{C}\delta^{-C_2\alpha_0},$$
$$\operatorname{supp} H_1 \subset \Big\{(w_1, \dots, w_m) \in \mathbb{R}^m : \min_{i=1,\dots,m} |w_i| \le e^{-M}\Big\}.$$

Indeed, let $\widetilde{F}_1 = C\delta^{-\alpha_0}\phi$ then $|F_1| \le \widetilde{F}_1$ and $\|\nabla^a \widetilde{F}_1\| \le \widetilde{C}\delta^{-C_2\alpha_0}$ since $\|\bar{F}\|_\infty \le C\delta^{-\alpha_0}$. Then let

$$H_1(w_1, \dots, w_m) = \widetilde{F}_1\big(\log|w_1|, \dots, \log|w_m|\big).$$

Since $\widetilde{F}_1$ is constant on Ω_2^c, H_1 is smooth. We have $\|H_1\| \le \widetilde{C}\delta^{-\alpha_0}$ and for all $a \ge 1$, $\nabla^a H_1 = 0$ on $(\log|w_1|, \dots, \log|w_m|) \in \operatorname{Int}(\Omega_2^c) \cup \operatorname{Int}(\Omega_1^c)$. In the remaining domain $(\log|w_1|, \dots, \log|w_m|) \in \bar{\Omega}_2 \cap \bar{\Omega}_1$, we have

$$\left|\frac{\partial H_1}{\partial w_1}\right| = \left|\frac{\partial \widetilde{F}_1}{\partial w_1}\frac{1}{|w_1|}\right| \le \widetilde{C}\delta^{-\alpha_0}\left|\frac{\partial\phi}{\partial w_1}\right|\frac{1}{|w_1|} \le \widetilde{C}\delta^{-C_2\alpha_0}\frac{1}{|w_1|},$$

where our constant C_2 can, as always, change from one line to another. Since $\log|w_1| \ge -M - 4$, $|w_1| \ge e^{-M-4}$. Thus, $|\frac{\partial H_1}{\partial w_1}| \le \widetilde{C}\delta^{-C_2\alpha_0}e^M$. Similarly for higher derivatives, we get that $\|\nabla^a H_1\| \le \widetilde{C}\delta^{-C_2\alpha_0}e^{3M}$. Choose $M = \log(\delta^{-3\alpha_0})$ then $\|\nabla^a H_1\| \le \widetilde{C}\delta^{-C_2\alpha_0}$ for all $0 \le a \le 3$. Applying Lemma A.1 to α_1 and $C_2\alpha_0$,

$$\begin{aligned}\mathbf{E}\left|F_1\left(\log\frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots\right)\right| &\le \mathbf{E}H_1\left(\frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots\right)\\ &\le \mathbf{E}H_1\left(\frac{|\widetilde{P}(z_1)|}{\sqrt{V(z_1)}}, \dots\right) + \widetilde{C}\delta^{C_2\alpha_0}.\end{aligned}$$

Since $H_1 = 0$ if $(\log|w_1|, \dots, \log|w_m|) \notin \Omega_1$, one has

$$\begin{aligned}&\mathbf{E}H_1\left(\frac{|\widetilde{P}(z_1)|}{\sqrt{V(z_1)}}, \dots, \frac{|\widetilde{P}(z_m)|}{\sqrt{V(z_m)}}\right)\\ &\quad\le \widetilde{C}\delta^{-\alpha_0}\mathbf{P}\left(\exists i \in \{1, \dots, m\} : \frac{|\widetilde{P}(z_i)|}{\sqrt{V(z_i)}} \le e^{-M} = \delta^{3\alpha_0}\right)\\ &\quad\le \widetilde{C}\delta^{-\alpha_0}m\delta^{3\alpha_0} \le \widetilde{C}\delta^{\alpha_0}.\end{aligned}$$

Thus, $\mathbf{E}|F_1(\log \frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots, \log \frac{|P(z_m)|}{\sqrt{V(z_m)}})| \le \widetilde{C}\delta^{\alpha_0}$. Finally, we will show that

$$\left|\mathbf{E}F_2\left(\log \frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots\right) - \mathbf{E}F_2\left(\log \frac{|\widetilde{P}(z_1)|}{\sqrt{V(z_1)}}, \dots\right)\right| \le \widetilde{C}\delta^{\alpha_0}.$$

Define $H_2 : \mathbb{R}^m \to \mathbb{R}$ by $H_2(w_1, \dots, w_m) = F_2(\log|w_1|, \dots, \log|w_2|)$. Since $\operatorname{supp} F_2 \subset \bar{\Omega}_2$, it follows that $\operatorname{supp}(H_2)$ is contained inside

$$\{(w_1, \dots, w_m) : \log|w_i| \ge -M - 4\ \forall i\} = \{(w_1, \dots, w_m) : |w_i| \ge \widetilde{C}\delta^{3\alpha_0}\ \forall i\}.$$

Thus, H_2 is well defined and smooth on $\mathbb{R}^m$. By a similar argument to the part about H_1, $\|\nabla^a H_2\| \le \widetilde{C}\delta^{-C_2\alpha_0}$ for all $0 \le a \le 3$. We can increase C_2 to have $C_2 \ge 1$. Applying Lemma A.1 to α_1 and $C_2\alpha_0$ gives

$$\begin{aligned}
&\left|\mathbf{E}F_2\left(\log \frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots\right) - \mathbf{E}F_2\left(\log \frac{|\widetilde{P}(z_1)|}{\sqrt{V(z_1)}}, \dots\right)\right| \\
&\quad = \left|\mathbf{E}H_2\left(\frac{|P(z_1)|}{\sqrt{V(z_1)}}, \dots, \frac{|P(z_m)|}{\sqrt{V(z_m)}}\right) - \mathbf{E}H_2\left(\frac{|\widetilde{P}(z_1)|}{\sqrt{V(z_1)}}, \dots, \frac{|\widetilde{P}(z_m)|}{\sqrt{V(z_m)}}\right)\right| \\
&\quad \le \widetilde{C}\delta^{C_2\alpha_0} \le \widetilde{C}\delta^{\alpha_0}.
\end{aligned}$$

This completes the proof. □

Acknowledgments. Part of this work was done at VIASM (Hanoi), and the authors would like to thank the institute for the support and hospitality.

REFERENCES

[1] ALON, N. and SPENCER, J. H. (2008). *The Probabilistic Method*, 3rd ed. Wiley, Hoboken, NJ. MR2437651

[2] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

[3] BHARUCHA-REID, A. T. and SAMBANDHAM, M. (1986). *Random Polynomials*. Academic Press, Orlando, FL. MR0856019

[4] BLEHER, P. and DI, X. (2004). Correlations between zeros of non-Gaussian random polynomials. *Int. Math. Res. Not. IMRN* **2004** 2443–2484. MR2078308

[5] BLOCH, A. and PÓLYA, G. (1931). On the Roots of Certain Algebraic Equations. *Proc. Lond. Math. Soc.* (2) **33** 102–114. MR1576817

[6] DAS, M. (1969). The average number of maxima of a random algebraic curve. *Proc. Cambridge Philos. Soc.* **65** 741–753. MR0239669

[7] DAS, M. (1972). Real zeros of a class of random algebraic polynomials. *J. Indian Math. Soc.* (*N.S.*) **36** 53–63. MR0322960

[8] DO, Y., NGUYEN, H. and VU, V. (2015). Real roots of random polynomials: Expectation and repulsion. *Proc. Lond. Math. Soc.* (3) **111** 1231–1260. MR3447793

[9] DUBHASHI, D. P. and PANCONESI, A. (2009). *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge Univ. Press, Cambridge. MR2547432

[10] EDELMAN, A. and KOSTLAN, E. (1995). How many zeros of a random polynomial are real? *Bull. Amer. Math. Soc.* (*N.S.*) **32** 1–37. MR1290398

[11] ERDÖS, P. and OFFORD, A. C. (1956). On the number of real roots of a random algebraic equation. *Proc. Lond. Math. Soc.* (3) **6** 139–160. MR0073870

[12] FARAHMAND, K. (1998). *Topics in Random Polynomials. Pitman Research Notes in Mathematics Series* **393**. Longman, Harlow. MR1679392

[13] HOUGH, J. B., KRISHNAPUR, M., PERES, Y. and VIRÁG, B. (2009). *Zeros of Gaussian Analytic Functions and Determinantal Point Processes. University Lecture Series* **51**. Amer. Math. Soc., Providence, RI. MR2552864

[14] IBRAGIMOV, I. A. and MASLOVA, N. B. (1968). The average number of zeros of random polynomials. *Vestnik Leningrad. Univ.* **23** 171–172. MR0238376

[15] IBRAGIMOV, I. A. and MASLOVA, N. B. (1971). The average number of real roots of random polynomials. *Sov. Math., Dokl.* **12** 1004–1008. MR0292134

[16] IBRAGIMOV, I. A. and MASLOVA, N. B. (1971). On the expected number of real zeros of random polynomials. I. Coefficients with zero means. *Theory Probab. Appl.* **16** 228–248.

[17] IBRAGIMOV, I. A. and MASLOVA, N. B. (1971). On the expected number of real zeros of random polynomials. II. Coefficients with non-zero means. *Theory Probab. Appl.* **16** 485–493.

[18] KAC, M. (1943). On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.* **49** 314–320. MR0007812

[19] KAC, M. (1949). On the average number of real roots of a random algebraic equation. II. *Proc. Lond. Math. Soc.* (2) **50** 390–408. MR0030713

[20] LITTLEWOOD, J. E. and OFFORD, A. C. (1943). On the number of real roots of a random algebraic equation. III. *Rec. Math.* [*Mat. Sbornik*] *N.S.* **12** 277–286. MR0009656

[21] LITTLEWOOD, J. E. and OFFORD, A. C. (1945). On the distribution of the zeros and a-values of a random integral function. I. *J. Lond. Math. Soc.* **20** 130–136. MR0019123

[22] LITTLEWOOD, J. E. and OFFORD, A. C. (1948). On the distribution of zeros and a-values of a random integral function. II. *Ann. of Math.* (2) **49** 885–952. MR0029981

[23] LOGAN, B. F. and SHEPP, L. A. (1968). Real zeros of random polynomials. *Proc. Lond. Math. Soc.* (3) **18** 29–35. MR0234512

[24] LOGAN, B. F. and SHEPP, L. A. (1968). Real zeros of random polynomials. II. *Proc. Lond. Math. Soc.* (3) **18** 308–314. MR0234513

[25] MASLOVA, N. B. (1974). On the variance of the number of real roots of random polynomials. *Theory Probab. Appl.* **19** 35–52.

[26] MASLOVA, N. B. (1975). On the distribution of the number of real roots of random polynomials. *Theory Probab. Appl.* **19** 461–473. MR0334327

[27] NAZAROV, F., NISHRY, A. and SODIN, M. (2014). Log-integrability of Rademacher Fourier series, with applications to random analytic functions. *St. Petersburg Math. J.* **25** 467–494. MR3184602

[28] NAZAROV, F. and SODIN, M. (2010). Random complex zeroes and random nodal lines. In *Proceedings of the International Congress of Mathematicians*, *Vol. III* 1450–1484. Hindustan Book Agency, New Delhi. MR2827851

[29] NAZAROV, F. and SODIN, M. (2012). Correlation functions for random complex zeroes: Strong clustering and local universality. *Comm. Math. Phys.* **310** 75–98. MR2885614

[30] NGUYEN, H., NGUYEN, O. and VU, V. (2016). On the number of real roots of random polynomials. *Commun. Contemp. Math.* **18** 1550052, 17. MR3493213

[31] NGUYEN, H. H. and VU, V. H. (2013). Small ball probability, inverse theorems, and applications. In *Erdös Centennial. Bolyai Soc. Math. Stud.* **25** 409–463. János Bolyai Math. Soc., Budapest. MR3203607

[32] PERES, Y. and VIRÁG, B. (2005). Zeros of the i.i.d. Gaussian power series: A conformally invariant determinantal process. *Acta Math.* **194** 1–35. MR2231337

[33] PRITSKER, I. E. and VARGA, R. S. (1999). Weighted rational approximation in the complex plane. *J. Math. Pures Appl.* (9) **78** 177–202. MR1677669

[34] RUDIN, W. (1986). *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York. MR0344043

[35] SAMBANDHAM, M. (1979). On the average number of real zeros of a class of random algebraic curves. *Pacific J. Math*. **81** 207–215. MR0543744

[36] SAMBANDHAM, M., GORE, H. and FARAHMAND, K. (1998). The average number of point [points] of inflection of random algebraic polynomials. *Stoch. Anal. Appl*. **16** 721–731. MR1632566

[37] STEVENS, D. C. (1969). The average number of real zeros of a random polynomial. *Comm. Pure Appl. Math*. **22** 457–477. MR0251003

[38] TAO, T. and VU, V. (2015). Local universality of zeroes of random polynomials. *Int. Math. Res. Not. IMRN* **2015** 5053–5139. MR3439098

[39] TAO, T. and VU, V. (2015). Random matrices: Universality of local spectral statistics of non-Hermitian matrices. *Ann. Probab.* **43** 782–874. MR3306005

[40] TODHUNTER, I. (2014). *A History of the Mathematical Theory of Probability*. Cambridge Univ. Press, Cambridge.

[41] WILKINS, J. E. JR. (1988). An asymptotic expansion for the expected number of real zeros of a random polynomial. *Proc. Amer. Math. Soc*. **103** 1249–1258. MR0955018

Y. DO
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22904
USA
E-MAIL: yendo@virginia.edu

O. NGUYEN
DEPARTMENT OF MATHEMATICS
YALE UNIVERSITY
NEW HAVEN, CONNECTICUT 06520
USA
E-MAIL: oanh.nguyen@yale.edu

V. VU
DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
LEVEL 4, BLOCK S17
10 LOWER KENT RIDGE ROAD
SINGAPORE 119076
E-MAIL: van.vu@yale.edu