RANDOM POLYNOMIALS: CENTRAL LIMIT THEOREMS FOR THE REAL ROOTS

OANH NGUYEN AND VAN VU

Abstract. The number of real roots has been a central subject in the theory of random polynomials and random functions since the fundamental papers of Littlewood-Offord and Kac in the 1940s.

In this paper, we establish the Central Limit Theorem for the number of real roots of random polynomials with coefficients having moderate growth, generalizing and strengthening a classical result of Maslova.

1. Introduction

Random polynomials, so simple to define but difficult to understand, have attracted generations of mathematicians. Typically, a random (algebraic) polynomial is defined as

$$P_n(x) := c_n x^n + \cdots + c_1 x + c_0,$$

where $\xi_i$ are iid copies of an (atom) random variable $\xi$ 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 distinct behaviors.

When $c_i = 1$ for all $i$, the polynomial $P_n$ is often referred to as the Kac polynomial. Even this special class provides great challenges, which have led to rich literature (see, for example, the books [3, 10] and the references therein).

Let $N_n(\mathbb{R})$ denote the number of real roots of $P_n$. A key problem in the theory of random polynomials is to understand the behavior of the random variable $N_n(\mathbb{R})$, with $n$ tending to infinity. As a matter of fact, this is the problem that started the whole field, with fundamental works of Littlewood-Offord [21, 22, 23] and Kac [19] from the 1940s.

The first natural question is to determine the expectation of $N_n(\mathbb{R})$. It took more than 20 years and the works of Kac [19], Erdős-Offord [9] and Ibragimov-Maslova [16, 17] to settle this problem for the Kac polynomial (the case $c_0 = \cdots = c_n = 1$). By now, the problem has been solved for many classes of random polynomials, with various choices for $c_i$ and under very general assumptions for $\xi_i$ (see the introduction of [27]; also [14] [8] [33] [31] [34] [37] [12] [18] [29] [30] [6] [35] [36] and the references therein).

The next, and perhaps more important, problem is to determine the variance and limiting distribution of $N_n(\mathbb{R})$. This problem is much harder and our understanding is far from complete. In the 1970s, Maslova [25] proved the Central Limit Theorem (CLT) for the Kac polynomial. Here and later, $\xrightarrow{\mathcal{D}}$ means convergence in distribution; $\mathcal{N}(0,1)$ denotes the standard normal distribution, $\mu_n := \mathbb{E} N_n(\mathbb{R}), \sigma_n := \sqrt{\text{Var} N_n(\mathbb{R})}$.

This work is partially supported by VIASM (Vietnam).
Theorem 1.1. [24, 25] Let $\varepsilon$ be a positive constant. Consider the Kac polynomial with the random variables $\xi_i$ being iid with mean zero, variance one, bounded $(2 + \varepsilon)$ moment, and $P(\xi_i = 0) = 0$. We have, as $n$ tends to infinity,

$$\frac{N_n(\mathbb{R}) - \mathbb{E}N_n(\mathbb{R})}{(\text{Var} \ N_n(\mathbb{R}))^{1/2}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Furthermore, $\text{Var} \ N_n(\mathbb{R}) = (K + o(1)) \log n$, where $K = \frac{4}{\pi} (1 - \frac{2}{\pi})$.

The proof of Maslova relied heavily on explicit computation that requires all the $c_i$ to be equal. Only very recently, Central Limit Theorems have been established for other classes of polynomials, via new methods. In 2015, Dalmao [5] established the CLT for binomial polynomials (the case when $c_i = \sqrt{\binom{n}{i}}$), and in 2018, Do and the second author [7] handled Weyl polynomials ($c_i = \frac{1}{\sqrt{i!}}$).

However, in both papers, the authors need to assume that the random variables $\xi_i$ are standard Gaussian and their arguments rely strongly on special properties of Gaussian processes. It remains a major challenge to extend these results to other random variables $\xi_i$ (Rademacher, for example).

For related results concerning random trigonometric polynomials, see [13, 2, 1].

The goal of this paper is to establish CLT for a large class of random polynomials where the deterministic coefficients $c_i$ grow polynomially. We will only need a mild assumption on the $\xi_i$, which is satisfied by most random variables used in practice. In fact, we can handle the more general setting which does not require the $\xi_i$ to be iid.

We consider

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

where $\xi_i$ are independent random variables and $c_i$ are deterministic coefficients satisfying the following conditions for some positive constants $N_0, \tau_1, \tau_2, \varepsilon$ and some constant $\rho > -1/2$.

(A1) The random variables $\xi_i$ are independent (but not necessarily identically distributed) real-valued random variables with unit variance and bounded $(2 + \varepsilon)$ moments, namely $\mathbb{E}|\xi_i|^{2 + \varepsilon} \leq \tau_2$.

(A2) $\mathbb{E}\xi_i = 0$ for all $i \geq N_0$.

(A3) The coefficients $c_i$ are deterministic real numbers that grow polynomially, namely $|c_i| \leq \tau_2$ for all $0 \leq i < N_0$,

and $\tau_1 i^\rho \leq |c_i| \leq \tau_2 i^\rho$ for all $N_0 \leq i < n$.

This class contains many interesting ensembles of polynomials including

- the Kac polynomial,
- all derivatives of the Kac polynomial (the zeroes of these polynomials are thus the critical points of the Kac polynomial),
- hyperbolic polynomials $P_n(x) = \sum_{i=0}^{n} \sqrt{\frac{L(L+1)\ldots(L+i-1)}{i!}} \xi_i x^i$ where $L$ is a positive constant (see [14, 6, 11] and the references therein).

Our main result is
Theorem 1.2. Assume that the polynomial \( P_n \) satisfies Conditions (A1)-(A3) and \( \text{Var} N_n(\mathbb{R}) = \Omega(\log n) \). Then \( \frac{N_n(\mathbb{R}) - \mu_n}{\sigma_n} \xrightarrow{d} N(0,1) \) where \( \mu_n = \mathbb{E} N_n(\mathbb{R}), \sigma_n = \sqrt{\text{Var} N_n(\mathbb{R})} \).

The condition \( \text{Var} N_n(\mathbb{R}) = \Omega(\log n) \) is guaranteed for all ensembles listed above thanks to the following lemma.

Lemma 1.3. Assume that the polynomial \( P_n \) satisfies Conditions (A1)-(A3) and there is a constant \( \varepsilon > 0 \) such that for all \( i \in \left[ n - n \exp\left(-\log^{5/6} n\right), n - \exp\left(\log^{1/5} n\right) \right) \),

\[
\frac{|c_i|}{|c_n|} = 1 + O\left(\exp\left(-\left(\log \log n\right)^{1+\varepsilon}\right)\right).
\]

Then \( \text{Var} N_n(\mathbb{R}) = \Omega(\log n) \).

The condition in this lemma is satisfied by all classes listed above. We obtain

Corollary 1.4. The CLT holds for the Kac polynomial and its derivatives. It also holds for hyperbolic polynomials.

Remark 1.5. When restricted to the Kac polynomial with \( \xi_i \) being iid copies of an atom variable \( \xi \), our result strengthens Maslova’s, as the condition \( P(\xi = 0) = 0 \) in Theorem 1.1 is removed.

Notation. We use standard asymptotic notations under the assumption that \( n \) tends to infinity. For two positive sequences \((a_n)\) and \((b_n)\), we say that \( a_n = \Omega(b_n) \) or \( b_n = O(a_n) \) or \( b_n \ll a_n \) if there exists a constant \( C > 0 \) such that \( b_n \leq C a_n \). If \( a_n \ll b_n \ll a_n \), we say that \( b_n = \Theta(a_n) \). If \( |c_n| \ll a_n \) for some sequence \((c_n)\), we also write \( c_n = O(a_n) \) or \( c_n \ll a_n \). If \( \lim_{n \to \infty} \frac{a_n}{b_n} = 0 \), we say that \( a_n = o(b_n) \).

2. The Universality Method

The key ingredient of our proof is the universality method. The general idea of this method is to show that limiting laws do not depend too much on the distribution of the atom variable \( \xi \) (or the variables \( \xi_i \) in general, if they are not iid). Once universality has been established, then it suffices to prove the desired law for the case in which the \( \xi_i \) are Gaussian, and here one can bring extra powerful tools such as properties of Gaussian processes; see [14, 8, 33, 31, 37, 12, 34, 15, 18, 29, 30].

The universality method has been powerful in studying local statistics such as the density or correlation functions concerning the number of roots in a small region (where the expectation is of order \( \Theta(1) \)) (see, for example, [37, 26, 6, 27]). In order to use it to prove the global law in this paper, we need to perform a number of considerably technical steps, linking local statistics to the global one. The proof for the Gaussian case itself also requires new ideas.

To study the real roots of \( P_n \), we divide the real line into two regions: a core region that contains most of the real roots and the remaining one that contains an insignificant number of real roots. Consider small numbers \( 0 \leq b_n < a_n < 1 \) that depend on \( n \) and satisfy the following property for all constant \( A > 0 \):

\[
a_n \ll \log^{-A} n.
\]

For example, \( a_n = \exp\left(-\left(\log n\right)^{1/5}\right) \). We define

\[
\mathfrak{J} := \mathfrak{J}_{a_n, b_n} := \pm(1 - a_n, 1 - b_n) \cup \pm(1 - a_n, 1 - b_n)^{-1}
\]
where for any given set $S$, we define $-S := \{ -x : x \in S \}$, $S^{-1} := \{ x^{-1} : x \in S \}$, and $\pm S := -S \cup S$. For appropriate choices of $a_n$ and $b_n$, this will be our core region.

For a subset $S \subset \mathbb{C}$, let $N_n(S) = N_{P_n}(S)$ be the number of roots of $P_n$ in $S$. Let $\xi_i$ be iid standard Gaussian random variables and set

$$\tilde{P}_n = \sum_{i=0}^{n} c_i \xi_i x^i.$$  

We denote by $\tilde{N}_n(S) = N_{\tilde{P}_n}(S)$ the number of zeros of $\tilde{P}_n$ in $S$.

Our main result on global universality of the real roots states that on the core $\mathcal{J}$, the distributions of the roots of $P_n$ and $\tilde{P}_n$ are approximately the same.

**Theorem 2.1.** Assume that the polynomial $P_n$ satisfies Conditions (A1)-(A3). There exist positive constants $C$ and $c$ such that for every $0 \leq b_n < a_n < 1$ satisfying (2), for sufficiently large $n$ and every function $F : \mathbb{R} \to \mathbb{R}$ whose derivatives up to order 3 are bounded by 1, we have

$$| \mathbb{E} F(N_n(\mathcal{J})) - \mathbb{E} F(\tilde{N}_n(\mathcal{J})) | \leq Ca_n^c + Cn^{-c}.$$  

Since $N_n(\mathcal{J})$ is always an integer, for every real number $a_0 \in \mathbb{R}$,

$$\mathbb{P}(N_n(\mathcal{J}) \leq a_0) = \mathbb{P}(N_n(\mathcal{J}) \leq \lfloor a_0 \rfloor) = \mathbb{E}(F(N_n(\mathcal{J})))$$

where $F$ is any smooth function that takes values in $[0,1]$ and $1_{(-\infty,|a_0|]} \leq F \leq 1_{(-\infty,|a_0|+1]}$. Therefore, Theorem 2.1 implies

$$| \mathbb{P}(N_n(\mathcal{J}) \leq a_0) - \mathbb{P}(\tilde{N}_n(\mathcal{J}) \leq a_0) | \leq Ca_n^c + Cn^{-c}. \tag{4}$$

Using Theorem 2.1 (not in the straightforward way), we deduce the following corollary

**Corollary 2.2.** Assume that the polynomial $P_n$ satisfies Conditions (A1)-(A3). Let $k \geq 1$ be an integer. There exist positive constants $C$ and $c$ such that for every $0 \leq b_n < a_n < 1$ satisfying (2) and for sufficiently large $n$, we have

$$| \mathbb{E} \left( N_n^k(\mathcal{J}) \right) - \mathbb{E} \left( \tilde{N}_n^k(\mathcal{J}) \right) | \leq Ca_n^c + Cn^{-c}.$$  

In particular,

$$| \text{Var} \left( N_n(\mathcal{J}) \right) - \text{Var} \left( \tilde{N}_n(\mathcal{J}) \right) | \leq Ca_n^c + Cn^{-c}.$$  

Next, we show that the contribution outside of the core is negligible.

**Proposition 2.3.** Assume that the polynomial $P_n$ satisfies Conditions (A1)-(A3). Let $k \geq 2$ be an integer. There exists a positive constant $C$ such that for every $0 \leq b_n < a_n < 1$ satisfying (2) and for sufficiently large $n$, we have

$$\mathbb{E} N_n^k(\mathbb{R} \setminus \mathcal{J}) \leq \begin{cases} C \left( \log a_n^{-1} \right)^{2k} + \log^k(nb_n) & \text{if } b_n \geq 1/n, \\ C \left( \log a_n^{-1} \right)^{2k} & \text{if } b_n < 1/n. \end{cases} \tag{5}$$

To prove Theorem 1.2 and Lemma 1.3, we use the universality results stated in Theorem 2.1, Corollary 2.2, and Proposition 2.3 to reduce to the Gaussian case (i.e., the case in which the $\xi_i$ are iid standard Gaussian) with roots restricted to the core $\mathcal{J}$. In particular, we prove...
Lemma 2.4. Assume that the polynomial $P_n$ satisfies Conditions (A1)-(A3). Let $c < 1$ be any positive constant, then for any $a_n, b_n$ satisfying

\[(6) \quad (\log n)^2/n \leq b_n < a_n \leq \exp(- (\log n)^{c}), \quad \log a_n = \Theta(\log n), \quad \text{and} \quad \text{Var} \tilde{N}_n(\mathcal{J}) = \Omega(\log n),\]

we have\[
\frac{\tilde{N}_n(\mathcal{J}) - \mathbb{E}\tilde{N}_n(\mathcal{J})}{\sqrt{\text{Var} \tilde{N}_n(\mathcal{J})}} \xrightarrow{d} \mathcal{N}(0, 1).
\]

Lemma 2.5. Under the assumptions of Lemma 1.3, we have\[
\text{Var} \tilde{N}_n(\mathbb{R}) = \Omega (\log n).
\]

To illustrate the method of universality, we include here the short proof of Theorem 1.2 and Lemma 1.3 assuming Lemma 2.4 and Lemma 2.5.

Proof of Lemma 1.3. We first choose $a_n$ and $b_n$ that satisfy all the conditions in Corollary 2.2 and make the right-hand side of (5) as small as $o(\log n)$ when $k = 2$. In particular, we let\[
a_n = \exp\left(- \log^{1/5} n\right), \quad b_n = \frac{1}{n a_n},\]

and\[
\mathcal{J} = \pm (1 - a_n, 1 - b_n) \cup \pm (1 - a_n, 1 - b_n)^{-1}.
\]

By the triangle inequality on the 2-norm, we obtain\[
\begin{align*}
\sqrt{\text{Var} N_n(\mathbb{R})} - \sqrt{\text{Var} N_n(\mathcal{J})} & \leq \sqrt{\text{Var} N_n(\mathbb{R} \setminus \mathcal{J})} \\
& \leq \sqrt{\mathbb{E}N_n^2(\mathbb{R} \setminus \mathcal{J})} = o\left(\sqrt{\log n}\right)
\end{align*}
\]

where in the last equation, we used Proposition 2.3. Since $\tilde{P}_n$ is just a special case of $P_n$ (where the random variables $\xi_i$ are iid Gaussian), we also have\[
\left|\sqrt{\text{Var} \tilde{N}_n(\mathbb{R})} - \sqrt{\text{Var} \tilde{N}_n(\mathcal{J})}\right| = o\left(\sqrt{\log n}\right).
\]

Combining this with Lemma 2.5, we obtain\[
\begin{align*}
\sqrt{\text{Var} \tilde{N}_n(\mathcal{J})} &= \sqrt{\text{Var} \tilde{N}_n(\mathbb{R})} + o\left(\sqrt{\log n}\right) \\
&= \Omega\left(\sqrt{\log n}\right).
\end{align*}
\]

Applying Corollary 2.2 and (8) yields\[
\text{Var} N_n(\mathcal{J}) = \text{Var} \tilde{N}_n(\mathcal{J}) + O(a_n^c) = \text{Var} \tilde{N}_n(\mathcal{J}) + o(\log n) = \Omega (\log n).
\]

From this and (7),\[
\sqrt{\text{Var} N_n(\mathbb{R})} = \sqrt{\text{Var} \tilde{N}_n(\mathcal{J})} + o\left(\sqrt{\log n}\right) = \Omega\left(\sqrt{\log n}\right).
\]

This completes the proof. \qed

Proof of Theorem 1.2. Let $a_n, b_n$ and $\mathcal{J}$ be as in the proof of Lemma 1.3. By the assumption that $\sigma_n = \sqrt{\text{Var} N_n(\mathbb{R})} \gg \sqrt{\log n}$ and by (7), we have\[
\sqrt{\text{Var} N_n(\mathcal{J})} = \sigma_n (1 + o(1)) \gg \sqrt{\log n}.
\]

By this and Corollary 2.2, we also have\[
\begin{align*}
\sqrt{\text{Var} \tilde{N}_n(\mathcal{J})} = \sqrt{\text{Var} \tilde{N}_n(\mathcal{J})} + o(1) &= \sigma_n (1 + o(1)) \gg \sqrt{\log n}.
\end{align*}
\]
Thus, (6) holds and so we can apply Lemma 2.4 to get
\[ \frac{\tilde{N}_n(J) - \mathbb{E}\tilde{N}_n(J)}{\sqrt{\text{Var} \, \tilde{N}_n(J)}} \xrightarrow{d} \mathcal{N}(0, 1). \]
Hence,
\[ \frac{N_n(J) - \mathbb{E}N_n(J)}{\sqrt{\text{Var} \, N_n(J)}} \xrightarrow{d} \mathcal{N}(0, 1) \]
because by (4), for any fixed \( a \in \mathbb{R} \),
\[ \mathbb{P} \left( \frac{N_n(J) - \mathbb{E}N_n(J)}{\sqrt{\text{Var} \, N_n(J)}} \leq a \right) = \mathbb{P} \left( \frac{\tilde{N}_n(J) - \mathbb{E}\tilde{N}_n(J)}{\sqrt{\text{Var} \, \tilde{N}_n(J)}} \leq a \right) + o(1) \xrightarrow{n \to \infty} \mathbb{P}(\mathcal{N}(0, 1) \leq a). \]
By Corollary 2.2,
\[ \mathbb{E}N_n(J) - \mathbb{E}\tilde{N}_n(J) = o(1). \]
Combining these with (9), we get
\[ \frac{N_n(J) - \mathbb{E}N_n(J)}{\sigma_n} \xrightarrow{d} \mathcal{N}(0, 1). \]
From Proposition 2.3, we have
\[ \mathbb{E}N_n(\mathbb{R} \setminus J) \ll \log^{2/5} n. \]
By Markov’s inequality, for any fixed \( a > 0 \), we have
\[ \mathbb{P} \left( \frac{N_n(\mathbb{R} \setminus J) - \mathbb{E}N_n(\mathbb{R} \setminus J)}{\sigma_n} \geq a \right) \leq \frac{1}{a\sigma_n} \mathbb{E} |N_n(\mathbb{R} \setminus J) - \mathbb{E}N_n(\mathbb{R} \setminus J)| \ll \frac{\log^{2/5} n}{a \log^{1/2} n} \xrightarrow{n \to \infty} 0. \]
Thus,
\[ \frac{N_n(\mathbb{R} \setminus J) - \mathbb{E}N_n(\mathbb{R} \setminus J)}{\sigma_n} \xrightarrow{d} 0 \]
Adding (10) and (11) completes the proof. \[ \square \]

In Section 7, we use universality again to prove Lemma 2.5. But in this case, we will reduce general coefficients \( c_i \) to the case when \( c_i = 1 \). In other words, we could swap random variables with different means or variances. This deviates significantly from standard swapping arguments that swap random variables with the same mean and variance.

The rest of the paper is organized as follows. Section 3 is devoted to the proof of Theorem 2.1, Section 4 for Corollary 2.2, Section 5 for Proposition 2.3, Section 6 for Lemma 2.4 and Section 7 for Lemma 2.5.

3. Proof of Theorem 2.1

To make the idea clearer, we first prove the theorem for \( N_n[1 - a_n, 1 - b_n] \) in place of \( N_n(J) \) and \( \tilde{N}_n[1 - a_n, 1 - b_n] \) in place of \( \tilde{N}_n(J) \). The original statement for \( N_n(J) \) and \( \tilde{N}_n(J) \) follows from the same arguments with some (merely technical) modifications explained in Remark 3.8.

Let \( \delta_i := a_n/2^i \) for \( i = 0, \ldots, M - 1 \) where \( M \) is the smallest number such that \( a_n/2^M \leq \max\{1/n, b_n\} \). Let \( \delta_M := \max\{1/n, b_n\} \). Note that \( M \ll \log n \). For each \( i \leq M - 1 \), let \( \tilde{N}_i \)
be the number of real roots of \( P_n \) in the interval \([1 - \delta_{i-1}, 1 - \delta_i]\). Let \( N_M \) be the number of real roots of \( P_n \) in the interval \([1 - \delta_{M-1}, 1 - b_n]\). We have, \( N_n(1 - a_n, 1 - b_n) = N_1 + \cdots + N_M \). \( \square \)

Theorem 2.1 is deduced from the following more general result that can be of independent interest.

**Proposition 3.1.** Let \( \hat{F} : \mathbb{R}^M \to \mathbb{R} \) be any function whose every partial derivative up to order 3 is bounded by 1. We have

\[
\left| \mathbb{E}\hat{F}(N_1, \ldots, N_M) - \mathbb{E}\hat{F} \left( \hat{N}_1, \ldots, \hat{N}_M \right) \right| = O(\delta_0^3).
\]

To deduce Theorem 2.1, let \( \hat{F} \) be the function defined by \( \hat{F}(x_1, \ldots, x_M) = F(x_1 + \cdots + x_M) \). It is easy to check that \( \left\| \partial^{(3)} \hat{F} \right\|_\infty \leq 1 \) where \( \max_{0 \leq |\alpha| \leq 3} \left\| \partial^\alpha \hat{F} \right\|_\infty \) being the supremum of all partial derivatives up to order 3 of \( F \). By applying Proposition 3.1 to this \( \hat{F} \), Theorem 2.1 follows.

**Proof of Proposition 3.1.** Let \( \alpha \) be a sufficiently small positive constant. Let \( \varphi_0 \) be a smooth function taking values in \([0, 1]\), supported on \([-1, 1]\) and equal 1 at 0 with \( \left\| \varphi_0^{(a)} \right\|_\infty = O(1) \) for all \( 0 \leq a \leq 3 \). For \( 1 \leq i \leq M \), let \( \varphi_i \) be a smooth function taking values in \([0, 1]\), supported on \([1 - \delta_{i-1} - \delta_i^{1+\alpha}, 1 - \delta_i + \delta_i^{1+\alpha}] \) and equal 1 on \([1 - \delta_{i-1}, 1 - \delta_i]\) with \( \left\| \varphi_i^{(a)} \right\|_\infty = O\left(\delta_i^{-a(1+\alpha)}\right) \) for all \( 0 \leq a \leq 3 \).

We approximate the indicator of the interval \([1 - \delta_{i-1}, 1 - \delta_i]\) by the following function defined on the complex plane

\[
\varphi_i(z) := \varphi_i(\text{Re}(z))\varphi_0 \left( \frac{\text{Im}(z)}{\delta_i^{1+\alpha}} \right).
\]

We have for all \( 0 \leq a \leq 3 \),

\[
\left\| \partial^{(a)} \varphi_i \right\|_\infty = O\left(\delta_i^{-a(1+\alpha)}\right).
\]

Let \((\zeta_j)_{j=1}^n\) be the roots of \( P_n \). We'll prove later the following lemma which asserts that the \( N_i \) (as a sum of indicator functions) are approximated very well by \( \sum_{j=1}^n \varphi_i(\zeta_j) \). The fact that the functions \( \varphi_i \) are smooth allows us to apply analytical tools.

**Lemma 3.2.** We have

\[
\mathbb{E}\hat{F}(N_1, \ldots, N_M) - \mathbb{E}\hat{F} \left( \sum_{j=1}^n \varphi_1(\zeta_j), \ldots, \sum_{j=1}^n \varphi_M(\zeta_j) \right) \ll \delta_0^{3/8}.
\]

Assuming Lemma 3.2 it remains to show that

\[
\mathbb{E}\hat{F} \left( \sum_{j=1}^n \varphi_1(\zeta_j), \ldots, \sum_{j=1}^n \varphi_M(\zeta_j) \right) = \mathbb{E}\hat{F} \left( \sum_{j=1}^n \varphi_1(\tilde{\zeta}_j), \ldots, \sum_{j=1}^n \varphi_M(\tilde{\zeta}_j) \right) + O(\delta_0^3)
\]

where \( \tilde{\zeta}_j \) are the roots of \( \tilde{P}_n \).

---

1 If \( a_n \leq 1/n \), we set \( M = 1, \delta_0 = \delta_1 = 1/n \) and \( N = N_1 \) to be the number of real roots of \( P_n \) in the interval \([1 - a_n, 1 - b_n]\).

Generally, there is no difference in our proof if an interval of interest includes one of its endpoints or not. So, for example, if one cares about \( N_n(1 - a_n, 1 - b_n) \) instead of \( N_n(1 - a_n, 1 - b_n) \), one can use the exact same analysis.
Lemma 3.3. We will show the following lemma:

\begin{equation}
\sum_{j=1}^{n} \varphi_i(\zeta_j) - \frac{2\delta_i^2}{9m_i} \sum_{k=1}^{m_i} \log |P_n(w_{ik})| \vartriangle \varphi_i(w_{ik}) = O(\delta_i^5)
\end{equation}

with probability at least \(1 - O(\delta_i^5)\), where \(w_{ik}\) are chosen independently, uniformly at random from the ball \(B(1 - 3\delta_i/2, 2\delta_i/3)\) and are independent of all previous random variables. We defer the proof of this fact to Appendix 8.1 as it is similar to proving [6, Equation (4.20)].

Since \(\sum_{i=1}^{M} \delta_i^5 \ll \delta_0^5\), by applying (15), the left-hand side of (14) equals

\[ E K(\log |P_n(w_{ik})|)_{i=1,\ldots,M} + O(\delta_0^5) \]

and the right-hand side of (14) equals

\[ E K(\log \hat{P}_n(w_{ik})|)_{i=1,\ldots,M} + O(\delta_0^5) \]

where

\[ K(x_{ik})_{i=1,\ldots,M} := \hat{F}\left(\frac{2\delta_i^2}{9m_1} \sum_{k=1}^{m_1} x_{1k} \vartriangle \varphi_1(w_{1k}), \ldots, \frac{2\delta_i^2}{9m_M} \sum_{k=1}^{m_M} x_{Mk} \vartriangle \varphi_M(w_{Mk})\right). \]

By (12) and the assumption on the derivatives of \(\hat{F}\),

\begin{equation}
\|K\|_{\infty} = O(1), \quad \left\| \frac{\partial K}{\partial x_{ik}} \right\|_{\infty} = O(\delta_i^{-2\alpha}), \quad \left\| \frac{\partial^2 K}{\partial x_{ik} \partial x'_{i'k'}} \right\|_{\infty} = O(\delta_i^{-2\alpha} \delta_{i'}^{-2\alpha}),
\end{equation}

and

\begin{equation}
\left\| \frac{\partial^3 K}{\partial x_{ik} \partial x'_{i'k'} \partial x''_{i''k''}} \right\|_{\infty} = O(\delta_i^{-2\alpha} \delta_{i'}^{-2\alpha} \delta_{i''}^{-2\alpha})
\end{equation}

for all \(i, i', i'', k, k', k''\).

We will show the following lemma.

Lemma 3.3. There exists a constant \(\alpha_0 > 0\) such that for every constant \(\alpha \in (0, \alpha_0]\), every function \(K : \mathbb{R}^{m_1 + \cdots + m_M} \to \mathbb{R}\) that satisfies (16) and every \(w_{ik}\) in \(B(1 - 3\delta_i/2, 2\delta_i/3)\), we have

\[ |E K(\log |P_n(w_{ik})|)_{i=1,\ldots,M} - E K(\log |\hat{P}_n(w_{ik})|)_{i=1,\ldots,M}| = O(\delta_0^5). \]

By assuming Lemma 3.3 and conditioning on the \(w_{ik}\), (14) follows. So does Proposition 3.1.

Proof of Lemma 3.3. By the derivative assumption on \(\hat{F}\), we have

\[ \hat{F}(N_1, \ldots, N_M) - \hat{F}\left(\sum_{j=1}^{n} \varphi_1(\zeta_j), \ldots, \sum_{j=1}^{n} \varphi_M(\zeta_j)\right) \ll \sum_{i=1}^{M} |N_i - \sum_{j=1}^{n} \varphi_i(\zeta_j)|. \]

For each \(i \leq M, |N_i - \sum_{j=1}^{n} \varphi_i(\zeta_j)|\) is bounded by the number of roots of \(P_n\) in the union of the sets \(S_1, S_2, S_3\) where \(S_1\) is the set of all complex numbers whose real part lies in \([1 - \delta_{i-1} - \delta_i^{1+\alpha}, 1 - \delta_i + \delta_i^{1+\alpha}] \setminus \{0\}\) and complex part in \([-\delta_i^{1+\alpha}, \delta_i^{1+\alpha}]\), \(S_2 := [1 - \delta_{i-1} - \delta_i^{1+\alpha}, 1 - \delta_{i-1}]\), and \(S_3 := [1 - \delta_{i-1}, 1 - \delta_i + \delta_i^{1+\alpha}]\).

To show that the number of roots on these sets is negligible, we use the following lemma from [6].
Lemma 3.4. [6] Lemma 5.1] There exists a constant $\alpha_0 > 0$ such that for all $0 < \alpha \leq \alpha_0$ and all $x \in \mathbb{R}$ with $|x| \in [1 - \delta_{i-1} - \delta_i^{1+\alpha}, 1 - \delta_i + \delta_i^{1+\alpha}]$,

$$P(N_n(B(x, 2\delta_i^{1+\alpha})) \geq 2) = O_\alpha(\delta_i^{3\alpha/2})$$

where $B(x, R)$ is the disk with center $x$ and radius $R$ in the complex plane.

To show that $E N_n(S_1)$ is small, we note that $S_1$ is contained in a union of $\Theta(\delta_i^{-\alpha})$ small balls of radius $2\delta_i^{1+\alpha}$. By Lemma 3.4, the union bound and the fact that the complex roots come in conjugate pairs, we have

$$P(N_n(S_1) > 0) \leq \sum_{\text{small balls}} P(\text{number of roots in a small ball is at least } 2) \ll \delta_i^{-\alpha} \delta_i^{3\alpha/2} = \delta_i^{\alpha/2}.$$

Thus, $N_n(S_1) = 0$ except on an event, named $A_1$, of probability at most $O(\delta_i^{\alpha/2})$.

To show that the contribution of $A_1$ is negligible and to conclude that $E N_n(S_1) = O(\delta_i^{\alpha/8})$, we use the following lemma on large deviation of the number of roots.

Lemma 3.5 (Bounded number of roots). For any positive constants $A$ and $k$, there exists a constant $C$ such that for every $n \geq C$, every $1/n \leq \delta \leq 1/C$ and $z \in \mathbb{C}$ with $1 - 2\delta \leq |z| \leq 1 - \delta + 1/n$, we have

$$P(N_n(B(z, \delta/2)) \geq C \log(1/\delta)) \leq C \delta^A,$$

and

$$E N_n^k(B(z, \delta/2)) \leq C \log^k(1/\delta).$$

Assume Lemma 3.5. Since $S_1 \subseteq B\left(1 - \frac{3\delta_i}{2}, \frac{\delta_i}{2} + \delta_i^{1+\alpha}\right)$, applying (18) to this ball and Hölder’s inequality, we obtain

$$E N_n(S_1) = E N_n(S_1) 1_{A_1} \leq (E N_n^2(S_1))^{1/2} (P(A_1))^{1/2} \leq C \delta_i^{\alpha/4} \log \frac{1}{\delta_i} \leq C \delta_i^{\alpha/8}.$$

For $S_2 \cup S_3$, by [6] Theorem 2.4, we have $E N_n(S_2 \cup S_3) = E \widetilde{N}_n(S_2 \cup S_3) + O(\delta_i^{\alpha/2})$. To estimate $E \widetilde{N}_n(S_2 \cup S_3)$, we use Kac-Rice formula ([19] [8]),

$$E \widetilde{N}_n(a, b) = \frac{1}{\pi} \int_a^b \sqrt{\sum_{i=0}^n \sum_{j=i+1}^n c_i^2 c_j^2 (j - i)^2 t^{2i + 2j - 2}} \ dt.$$

By a standard algebraic manipulation which is elaborated in Appendix 8.2, we get

$$E \widetilde{N}_n(S_2 \cup S_3) = O\left(\delta_i^{\alpha/2}\right).$$

Putting the bounds together, we obtain $E N_n(S_2 \cup S_3) = O(\delta_i^{\alpha/2})$.

By combining this with (19), it follows that the left-hand side of (13) is bounded by $O\left(\sum_{i=1}^M \delta_i^{\alpha/8}\right) = O\left(\delta_0^{\alpha/8}\right)$, proving (13) and Lemma 3.2. \qed
Proof of Lemma 3.5. We use Jensen’s inequality (whose proof can be found in, for example, [27, Appendix 15.5]) which asserts that for every entire function $f$, every $z \in \mathbb{C}$ and $0 < r < R$,

\[(22) \quad N_f (B(z, r)) \leq \frac{\log \frac{M_1}{M_2}}{\log \frac{R^2 + r^2}{2R}} \]

where $M_1 = \sup_{w \in B(z, R)} |f(w)|$ and $M_2 = \sup_{w \in B(z, r)} |f(w)|$. Applying this inequality to the polynomial $P_n$ gives

\[(23) \quad N_n (B(z, \delta/2)) \ll \log \frac{M_1}{M_2} \]

where $M_1 = \sup_{w \in B(z, 2\delta/3)} |P_n(w)|$ and $M_2 = \sup_{w \in B(z, \delta/2)} |P_n(w)|$.

We shall prove that for a large constant $C$ and for every $a \in [1, n\delta]$,

\[(24) \quad \mathbb{P} (N_n (B(z, \delta/2)) \geq C a - C \log \delta) \ll a^{-A} \delta^A \]

where the implicit constant only depends on $A$ and $C$. Setting $a = 1$, we obtain (17). Setting $A = 2k$, letting $a$ run from 1 to $n\delta$ and using the fact that $N_n(B(z, \delta/2)) \leq n$ with probability 1, we obtain (18), completing the proof.

From (23), to prove (24), if suffices to show that

\[(25) \quad \mathbb{P} (M_1 \geq \exp (Ca - C \log \delta)) \ll a^{-A} \delta^A \]

and

\[(26) \quad \mathbb{P} (M_2 \leq \exp (-Ca + C \log \delta)) \ll a^{-A} \delta^A. \]

Since

\[ M_1 \leq \sum_{i=0}^{n} |c_i||\xi_i||z|^i, \]

we have $\mathbb{E}M_1 \ll \delta^{-O(1)}$ by Conditions (A1) and (A3). The bound (25) then follows from Markov’s inequality.

For (26), writing $z = re^{i\theta}$ and observing that the set $\{ w = re^{i\theta'} : \theta' \in [\theta - \delta/10, \theta + \delta/10] \}$ is a subset of $B(z, \delta/2)$, we have

\[ \mathbb{P} (M_2 \leq \exp (-Ca + C \log \delta)) \leq \mathbb{P} \left( \sup_{\theta' \in [\theta - \delta/10, \theta + \delta/10]} \left| \sum_{j=0}^{n} c_j \xi_j r^j e^{ij\theta'} \right| \leq \exp (-Ca + C \log \delta) \right). \]

By taking the supremum outside, the right-hand side is at most

\[ \sup_{\theta' \in [\theta - \delta/10, \theta + \delta/10]} \mathbb{P} \left( \left| \sum_{j=0}^{n} c_j \xi_j r^j e^{ij\theta'} \right| \leq \exp (-Ca + C \log \delta) \right) \]

and hence, by projecting onto the real line and conditioning on the random variables $(\xi_j)_{j \notin [1, a/\delta]}$, it is bounded by

\[ \sup_{\theta' \in [\theta - \delta/10, \theta + \delta/10]} \sup_{Z \in \mathbb{R}} \mathbb{P} \left( \left| \sum_{j=1}^{a/\delta} c_j \xi_j r^j \cos(j\theta') - Z \right| \leq \exp (-Ca + C \log \delta) \right). \]

We use the following anti-concentration lemma from [27].
Lemma 3.6. [27] Lemma 9.2] Let \( \mathcal{E} \) be an index set of size \( M \in \mathbb{N} \), and let \( (\xi_j)_{j \in \mathcal{E}} \) be independent random variables satisfying Condition (A1). Let \( (e_j)_{j \in \mathcal{E}} \) be deterministic (real or complex) coefficients with \( |e_j| \geq \bar{c} \) for all \( j \) and for some number \( \bar{c} \in \mathbb{R}_+ \). Then for any constant \( B \geq 1 \), any interval \( I \subset \mathbb{R} \) of length at least \( M^{-B} \), there exists \( \theta' \in I \) such that

\[
\sup_{Z \in \mathbb{R}} P \left( \left| \sum_{j \in \mathcal{E}} e_j \xi_j \cos(j \theta') - Z \right| \leq \bar{c} M^{-16B^2} \right) \ll M^{-B/2}
\]

where the implicit constant depends only on \( B \) and the constants in Condition (A1).

Applying Lemma 3.6 with \( B = 2A, \mathcal{E} = [1, a/\delta], M = a/\delta, I = [\theta - \delta/10, \theta + \delta/10], e_j = c_j r^j \) and \( \bar{c} = \frac{\delta}{\sqrt{\log \delta}} \) (where we use Condition (A3) and the assumption that \( r = \sqrt{z} \geq 1 - 2\delta \) to get \( |e_j| \geq \bar{c} \) ), we obtain \( \theta' \in [\theta - \delta/10, \theta + \delta/10] \) such that for a sufficiently large constant \( C \),

\[
\sup_{Z \in \mathbb{R}} P \left( \left| \sum_{j = 1}^{a/\delta} c_j \xi_j r^j \cos(j \theta') - Z \right| \leq \exp \left( -C a + C^2 \log \delta \right) \right) \ll (a/\delta)^{-A} = a^{-A} \delta^A
\]

which gives (26) and completes the proof of Lemma 3.5. \( \square \)

In order to prove Lemma 3.3, we first prove the following smooth version where we disregard the singularity of the logarithm function. The proof of Lemma 3.3 follows by a routine smoothening argument that we defer to Appendix 8.3.

Lemma 3.7. There exists a constant \( c_0 > 0 \) such that for every \( \alpha \in (0, c_0] \), every smooth \(^2\) function \( L : \mathbb{C}^{m_1 + \cdots + m_M} \to \mathbb{R} \) that satisfies (16) and every \( w_{ik} \) in \( B(1 - 3\delta_i/2, 2\delta_i/3) \), we have

\[
\left| E L \left( \frac{P_n(w_{ik})}{\sqrt{V(w_{ik})}} \right)_{i=1,\ldots,M} - E L \left( \frac{\hat{P}_n(w_{ik})}{\sqrt{V(w_{ik})}} \right)_{i=1,\ldots,M} \right| = O(\delta_0^2),
\]

where \( V(w) := \sum_{j=N_0}^n |c_j|^2 |w|^{2j} \) and \( N_0 \) is the constant in Conditions (A2) and (A3).

Proof of Lemma 3.7. We use the Lindeberg swapping argument. Let \( P_{i_0}(z) = \sum_{i=0}^{i_0-1} c_i \xi_i z^i + \sum_{i=i_0}^n c_i \xi_i z^i \), for \( 0 \leq i_0 \leq n + 1 \). We have \( P_0 = P_n \) and \( P_{n+1} = \hat{P}_n \) and \( P_{i_0+1} \) is obtained from \( P_{i_0} \) by replacing the random variable \( \xi_i \) by \( \xi_{i_0} \). Let

\[
I_{i_0} := \left| E L \left( \frac{P_{i_0}(w_{ik})}{\sqrt{V(w_{ik})}} \right)_{ik} - E L \left( \frac{P_{i_0+1}(w_{ik})}{\sqrt{V(w_{ik})}} \right)_{ik} \right|.
\]

The left-hand side of (27) is bounded by \( \sum_{i_0=0}^n I_{i_0} \). Fix \( i_0 \in [N_0, n + 1] \) (where \( N_0 \) is the constant in Conditions (A2) and (A3)) and let

\[
Y_{ik} := \frac{P_{i_0}(w_{ik})}{\sqrt{V(w_{ik})}} - \frac{c_{i_0} \xi_{i_0} w_{ik}^{i_0}}{\sqrt{V(w_{ik})}}
\]

for \( 1 \leq i \leq M, 1 \leq k \leq m_i \). We have

\[
\frac{P_{i_0+1}(w_{ik})}{\sqrt{V(w_{ik})}} = Y_{ik} + \frac{c_{i_0} \xi_{i_0} w_{ik}^{i_0}}{\sqrt{V(w_{ik})}}.
\]

\(^2\)By “smooth”, we mean that \( L \) has continuous derivatives up to order 3.
Conditioned on the $\xi_j$ and $\tilde{\xi}_j$ for all $j \neq i_0$, the $Y_{ik}$ are fixed. To bound $I_{i_0}$, we reduce to bounding

\begin{equation}
\label{eq:di_0_def}
d_{i_0} := \left| \mathbb{E}_{\xi_{i_0}, \tilde{\xi}_{i_0}} \hat{L} \left( \frac{c_{i_0} \xi_{i_0} w_{ijk}}{\sqrt{V(w_{ik})}} \right)_{ik} - \mathbb{E}_{\xi_{i_0}, \tilde{\xi}_{i_0}} \hat{L} \left( \frac{c_{i_0} \tilde{\xi}_{i_0} w_{ijk}}{\sqrt{V(w_{ik})}} \right)_{ik} \right|,
\end{equation}

where $\hat{L} = \hat{L}_{i_0}(x_{ik})_{ik} := L(Y_{ik} + x_{ik})_{ik}$. Note that this function $\hat{L}$ also satisfies \ref{eq:2} because $L$ does.

Let $a_{ik,i_0} = \frac{c_{i_0} w_{ijk}^{i_0}}{\sqrt{V(w_{ik})}}$. By Condition \ref{eq:A3}, we have

\begin{equation}
\label{eq:V_bound}
V(w_{ik}) \gtrsim \sum_{j=\delta_i^{-1}}^{2\delta_i^{-1}} j^{2\rho} (1 - 13\delta_i/6)^{2j} \gtrsim \delta_i^{-1-2\rho}
\end{equation}

and

\begin{equation}
\label{eq:constraint_bound}
|c_{i_0} w_{ikk}^{i_0}| \ll i_0^\rho (1 - \delta_i/6)^{i_0} \ll i_0^\rho \exp(-i_0 \delta_i/6) \ll \max\{1, \delta_i^{-\rho}\}.
\end{equation}

Since $\rho > -1/2$, we have from \ref{eq:V_bound} and \ref{eq:constraint_bound} that

\begin{equation}
|a_{ik,i_0}| \ll \delta_i^{\alpha_1}
\end{equation}

for some constant $\alpha_1 > 0$. Taylor expanding $\hat{L}$ around the origin, we obtain

\begin{equation}
\hat{L}(a_{ik,i_0} \xi_{i_0})_{ik} = \hat{L}(0) + \hat{L}_1 + \text{err}_1,
\end{equation}

where

$$
\hat{L}_1 = \frac{d \hat{L}(a_{ik,i_0} \xi_{i_0})_{ik}}{dt} \bigg|_{t=0} = \sum_{ik} \frac{\partial \hat{L}(0)}{\partial \text{Re}(z_{ik})} \text{Re}(a_{ik,i_0} \xi_{i_0}) + \sum_{ik} \frac{\partial \hat{L}(0)}{\partial \text{Im}(z_{ik})} \text{Im}(a_{ik,i_0} \xi_{i_0}).
$$

Since $\hat{L}$ satisfies \ref{eq:2}, we have

$$
|\text{err}_1| \leq \sup_{t \in [0,1]} \left| \frac{d^2 \hat{L}(a_{ik,i_0} \xi_{i_0})_{ik}}{dt^2} \right| \leq |\xi_{i_0}|^2 \sum_{ik,i^\prime k^\prime} |a_{ik,i_0} a_{i^\prime k^\prime,i_0}| \delta_i^{-2\alpha} \delta_{i^\prime}^{-2\alpha} \ll |\xi_{i_0}|^2 \left( \sum_{ik} |a_{ik,i_0}| \delta_i^{-2\alpha} \right)^2.
$$

Expanding to the next derivative, we have, in a similar manner,

\begin{equation}
\hat{L}(a_{ik,i_0} \xi_{i_0})_{ik} = \hat{L}(0) + \frac{1}{2} \hat{L}_2 + \text{err}_2,
\end{equation}

where $\hat{L}_2 = \frac{d^2 \hat{L}(a_{ik,i_0} \xi_{i_0})_{ik}}{dt^2} \bigg|_{t=0}$ and

$$
|\text{err}_2| \ll |\xi_{i_0}|^3 \left( \sum_{ik} |a_{ik,i_0}| \delta_i^{-2\alpha} \right)^3.
$$

By definition, $|\text{err}_2| = \left| \text{err}_1 - \frac{1}{2} \hat{L}_2 \right| \ll |\xi_{i_0}|^2 \left( \sum_{ik} |a_{ik,i_0}| \delta_i^{-2\alpha} \right)^2$. Using interpolation, H"older’s inequality and $m_i = \delta_i^{-11\alpha}$, we get

$$
|\text{err}_2| \ll |\xi_{i_0}|^{2+\varepsilon} \left( \sum_{ik} |a_{ik,i_0}| \delta_i^{-2\alpha} \right)^{2+\varepsilon} \ll |\xi_{i_0}|^{2+\varepsilon} M^{1+\varepsilon} \sum_{i=1}^{M} \delta_i^{-50\alpha} \left( \sum_{k=1}^{m_i} |a_{ik,i_0}| \right)^{(2+\varepsilon)/2}.
$$
All of these estimates also hold for \( \tilde{\xi}_{i_0} \) in place of \( \xi_{i_0} \). Since \( \xi_{i_0} \) and \( \tilde{\xi}_{i_0} \) have the same first and second moments and they both have bounded \( (2 + \varepsilon) \) moments, we get

\[
d_{i_0} = |\mathbb{E} \text{err}_2| \ll M^{1+\varepsilon} \sum_{i_1=1}^{M} \delta_i^{-50\alpha} \left( \sum_{k=1}^{m_i} |a_{ik,i_0}|^2 \right)^{(2+\varepsilon)/2}.
\]

Taking expectation with respect to the remaining variables shows that the same upper bound holds for \( I_{i_0} \) for all \( N_0 \leq i_0 \leq n+1 \). By (31), choosing \( \alpha \) to be sufficiently small compared to \( \alpha_1 \), we have \( (\sum_{k=1}^{m_i} |a_{ik,i_0}|^2)^{\varepsilon/2} \ll \delta_i^{(12\alpha_1-11\alpha)\varepsilon/2} \ll \delta_i^{100\alpha} \). Hence,

\[
\sum_{i_0=N_0}^{n+1} I_{i_0} \ll M^{1+\varepsilon} \sum_{i_0=N_0}^{n+1} \sum_{i=1}^{M} \delta_i^{50\alpha} \sum_{k=1}^{m_i} |a_{ik,i_0}|^2 \ll \log^2 n \sum_{i=1}^{M} \delta_i^{2\alpha} \ll (\log^2 n) \delta_0^{2\alpha} \ll \delta_0^\alpha,
\]

where we used \( M \ll \log n \), \( \sum_{i_0=N_0}^{n+1} |a_{ik,i_0}|^2 = 1 \) and (3).

For \( 0 \leq i_0 < N_0 \), instead of (32) and (33), we use mean value theorem to get a rough bound

\[
\hat{L}(a_{ik,i_0} \xi_{i_0})_{ik} = \hat{L}(0) + O \left( \xi_{i_0} \sum_{i=1}^{M} \delta_i^{-2\alpha} |a_{ik,i_0}| \right),
\]

which by the same arguments as above gives

\[
I_{i_0} \ll M^{1/2} \left( \sum_{i=1}^{M} \delta_i^{-50\alpha} \sum_{k=1}^{m_i} |a_{ik,i_0}|^2 \right)^{1/2} \ll \log^{1/2} n \sum_{i=1}^{M} \delta_i^{2\alpha - 61\alpha} \ll \delta_0^\alpha.
\]

Taking all these bounds together, we get \( \sum_{i_0=0}^{n+1} I_{i_0} \ll \delta_0^\alpha \). This completes the proof of Lemma 3.7.

**Remark 3.8.** Going back to the deduction at the beginning of the proof, to prove the original statement of Theorem 2.1 for \( N_n(3) \) and \( \tilde{N}_n(3) \), at first, we decompose \( N_n(3) \) and \( \tilde{N}_n(3) \) into the sum of the numbers of real roots in the intervals \( [1-\delta_{i_1}, 1-\delta_i) \), \( [1-\delta_{i_1}, 1-\delta_i) \), \( [1-\delta_{i_1}, 1-\delta_i) \) and \( [1-\delta_{i_1}, 1-\delta_i) \). The intervals \( [1-\delta_{i_1}, 1-\delta_i) \) have been dealt with as we have seen. The number of real roots of \( P_n \) in the intervals \( [1-\delta_{i_1}, 1-\delta_i) \) is treated as the number of real roots in \( [1-\delta_{i_1}, 1-\delta_i) \) of the polynomial \( P_n(z) = \sum_{i=0}^{n} (-1)^i c_i \xi_i z^i \). The number of real roots of \( P_n \) in the intervals \( [1-\delta_{i_1}, 1-\delta_i) \) is treated as the number of real roots in \( [1-\delta_{i_1}, 1-\delta_i) \) of the polynomial \( P_n(z) = \sum_{i=0}^{n} (-1)^i c_i \xi_i z^i \). And the number of real roots of \( P_n \) in the intervals \( [1-\delta_{i_1}, 1-\delta_i) \) is treated as the number of real roots in \( [1-\delta_{i_1}, 1-\delta_i) \) of the polynomial \( R_n(z) \). Like \( P_n \), the polynomial \( P_n(z) \) satisfies Condition \( (A3) \) and hence, all of the above arguments work for \( P_n(-z) \). Observe that most of the coefficients \( \frac{c_i}{c_n} \xi_i z^i \) of \( R_n \) satisfy Condition \( (A3) \) with \( \rho = 0 \) (and the contribution of the remaining coefficients can be shown to be negligible). We leave it as an exercise for the interested reader that the arguments used for \( P_n \) also hold for \( R_n \) and hence also for \( R_n(-z) \).

### 4. Proof of Corollary 2.2

We define \( \delta_0, \ldots, \delta_M, N_1, \ldots, N_M \) as in the beginning of the proof of Theorem 2.1. Note that \( \delta_i \geq \delta_M \geq 1/n \) and \( \delta_0^\alpha = \Theta \left( \alpha_n^c + n^{-c} \right) \).
Assuming the first part of Corollary 2.2, the second part follows immediately by observing that

$$(\mathbb{E}N_n(\mathcal{J}))^2 - \left(\mathbb{E}\tilde{N}_n(\mathcal{J})\right)^2 \ll \delta_0^2 (2\mathbb{E}N_n(\mathcal{J}) + O(\delta_0^c)) \ll \delta_0^c n \ll \delta_0^{c/2}$$

where in the first inequality, we used the first part of Corollary 2.2 for $k = 1$, in the second inequality, we used (18) to get that

$$\mathbb{E}N_n(\mathcal{J}) \ll \sum_{i=1}^M \log(1/\delta_i) \leq \sum_{i=1}^M \log n \ll \log^2 n,$$

and in the last inequality, we used (2).

To prove the first part of Corollary 2.2, we first reduce to the interval $[1 - a_n, 1 - b_n]$ as explained in Remark 3.8, namely, it suffices to show that

$$(34) \quad \left| \mathbb{E}\left(N^k_n[1 - a_n, 1 - b_n]\right) - \mathbb{E}\left(\tilde{N}^k_n[1 - a_n, 1 - b_n]\right) \right| \leq C\delta_0^c.$$  

We write $N := N_n[1 - a_n, 1 - b_n]$, $\tilde{N} := \tilde{N}_n[1 - a_n, 1 - b_n]$. Let $A$ be the event on which $N \leq \log^4 n$ (here, 4 can be replaced by any large constant). Let $F$ be a smooth function that is supported on the interval $[-1, \log^4 n + 1]$ and $F(x) = x^k$ for all $x \in [0, \log^4 n]$. Since $N$ is always an integer, it holds that $N^k 1_A = F(N)$. The function $F$ can be chosen such that all of its derivatives up to order 3 are bounded by $O(\log^4 n)$. Applying Theorem 2.1 to the rescaled function $(\log n)^{-k} F$, we obtain

$$\left| \mathbb{E}[N^k 1_A] - \mathbb{E}[\tilde{N}^k 1_{\tilde{A}}] \right| = \left| \mathbb{E}[F(N)] - \mathbb{E}[F(\tilde{N})] \right| \ll \delta_0^c \log^{-4k} n \ll \delta_0^c$$

for some small constant $c$ where $\tilde{A}$ is the corresponding event on which $\tilde{N} \leq \log^4 n$.

To finish the proof, we show that the contribution from the complement of $A$ is negligible, i.e.,

$$\mathbb{E}[N^k 1_{A^c}] \ll \delta_0^c.$$  

Since $M \ll \log n \ll \delta_0^{-c/2}$ by (2) and since $N^k \leq M^k \sum_{i=0}^M N_i^k$, it suffices to show that for all $i$, $\mathbb{E}[N^i 1_{A^c}] \ll \delta_0^{2i}$. Let $\mathcal{A}_i$ be the event on which $N_i \leq \log^4 (1/\delta_i)$. Note that $\bigcap_{i=1}^M \mathcal{A}_i \subset A$. Let $A$ be a large constant. By (17) of Lemma 3.5, $P(\mathcal{A}_i^c) \ll \delta_i^A$. Thus,

$$P(\mathcal{A}) \leq \sum_{i=1}^M P(\mathcal{A}_i^c) \ll \sum_{i=1}^M \delta_i^A \ll \delta_0^A.$$  

This together with (18) of Lemma 3.5 give

$$\mathbb{E}[N^k 1_{A^c}] \ll \log^k n \sum_{i=1}^M \mathbb{E}[N^i 1_{A^c}] \ll \log^k n \sum_{i=1}^M \left(\mathbb{E}[N^i]^{1/2} \left(\mathbb{E}[\mathcal{A}^c]^{1/2} \ll \log^k n \sum_{i=1}^M \delta_i^{A/2} \left(\log 1/\delta_i\right)^{2k}.$$  

Since $\delta_i \geq 1/n$, the right most side is at most $(\log^4 n)\delta_0^{A/2} \ll \delta_0^{A/2 - 1} \ll \delta_0^c$ by (2) and by choosing $A \geq 3$. This completes the proof of Corollary 2.2.

5. Proof of Proposition 2.3

Let $\mathfrak{A}$ be the right-hand side of (5):

$$\mathfrak{A} := \begin{cases} 
(\log a_n^{-1})^{2k} + \log^k(nb_n) & \text{if } b_n \geq 1/n, \\
(\log a_n^{-1})^{2k} & \text{if } b_n < 1/n.
\end{cases}$$
Writing $\mathbb{R} \setminus \mathcal{I}$ as a union of four sets $T_1 := [0, 1] \setminus \mathcal{I}$, $T_2 := [-1, 0] \setminus \mathcal{I}$, $T_3 := (1, \infty) \setminus \mathcal{I}$ and $T_4 := (-\infty, -1) \setminus \mathcal{I}$ and using triangle inequality, we reduce Proposition 2.3 to showing that for each $1 \leq i \leq 4$,

\begin{equation}
\mathbb{E}N_n^k(T_i) = \mathbb{E}N_n^k(T_i) \ll \mathfrak{A}.
\end{equation}

As explained in Remark 3.8, it suffices to show (35) for $i = 1$. Since $\mathfrak{A} \gg 1$, by triangle inequality, (35) follows from showing that for some large constant $C$,

\begin{equation}
\mathbb{E}N_n^k[0, 1 - 1/C] \ll 1,
\end{equation}

and

\begin{equation}
\mathbb{E}N_n^k(1 - 1/C, 1 - a_n) \ll (\log a_n^{-1})^{2k},
\end{equation}

\begin{equation}
\mathbb{E}N_n^k(1 - a_n/n, 1) \ll 1,
\end{equation}

and

\begin{equation}
\mathbb{E}N_n^k(1 - b_n, 1 - a_n/n) \ll \mathfrak{A}
\end{equation}

where we note that if $1 - b_n > 1 - a_n/n$ then the interval $(1 - b_n, 1 - a_n/n)$ is empty and (39) is vacuously true.

The bound (36) is precisely the content of [4 Lemma 2.5].

For (37), dividing the interval $(1 - 1/C, 1 - a_n)$ into dyadic intervals $I_0 := (1 - 1/C, 1 - 1/2C]$, $I_1 := (1 - 2C, 1 - 1/2C]$, $I_2 := (1 - 2/C, 1 - a_n) - (1 - 2/C, 1 - 1/2C]$, $I_3 := (1 - 2/C, 1 - a_n)$ (where $\frac{1}{2C} \leq a_n > \frac{1}{2C + 1/C}$) and applying triangle inequality together with (18), we obtain

\begin{equation}
\left(\mathbb{E}N_n^k(1 - 1/C, 1 - a_n)\right)^{1/k} \leq \sum_{i=0}^m \left(\mathbb{E}N_n^k(I_i)\right)^{1/k} \ll \sum_{i=0}^m \log(2^i C) \ll (\log a_n^{-1})^2.
\end{equation}

Thus,

\begin{equation}
\mathbb{E}N_n^k(1 - 1/C, 1 - a_n) \ll (\log a_n^{-1})^{2k}
\end{equation}

proving (37).

To prove (38) and (39), applying (34) to the intervals $(1 - a_n/n, 1)$ and $(1 - b_n, 1 - a_n/n)$, we get

\begin{equation}
\left|\mathbb{E}N_n^k(1 - a_n/n, 1) - \tilde{\mathbb{E}}N_n^k(1 - a_n/n, 1)\right| \ll n^{-c} \ll 1
\end{equation}

and

\begin{equation}
\left|\mathbb{E}N_n^k(1 - b_n, 1 - a_n/n) - \tilde{\mathbb{E}}N_n^k(1 - b_n, 1 - a_n/n)\right| \leq Cb_n^c + Cn^{-c} \ll 1 \ll \mathfrak{A}.
\end{equation}

Thus, it remains to prove (38) and (39) when the random variables $\xi_i$ are iid standard Gaussian. So for the rest of this proof, we assume that it is the case. For (38), we use Hölder’s inequality and (18) to conclude that

\begin{equation}
\mathbb{E}N_n^k(1 - a_n/n, 1) \leq \left(\mathbb{E}N_n^{2k-1}(1 - a_n/n, 1)\right)^{1/2} \left(\mathbb{E}N_n(1 - a_n/n, 1)\right)^{1/2}
\end{equation}

\begin{equation}
\ll (\log n)^{2k-1} \left(\mathbb{E}N_n(1 - a_n/n, 1)\right)^{1/2}.
\end{equation}

Using the Kac-Rice formula (20), we get

\begin{equation}
\mathbb{E}N_n(1 - a_n/n, 1) = \frac{1}{\pi} \int_{1-\frac{a_n}{n}}^{1} \sqrt{\sum_{i=0}^n \sum_{j=i+1}^n c_i^2 c_j^2 (j - i)^2 l^{2i+2j-2}} \, dt \ll \int_{1-\frac{a_n}{n}}^{1} n dt = a_n.
\end{equation}
where we used $|j - i| \leq n + 1$ and $\sum_{i=0}^{n} \sum_{j=i+1}^{n} c_{j}^{2} t^{2i+2j} = (\sum_{i=0}^{n} c_{i}^{2} t^{2i})^{2}$. Plugging this into (40) and using (2), we obtain

$$\mathbb{E}N_{n}^{k} \left(1 - \frac{a_{n}}{n}, 1\right) \ll (\log n)^{2k-1} a_{n}^{1/2} \ll 1.$$  

Finally, we prove (39). For any $x, y \in \mathbb{R}$, let

$$V(x) := \text{Var } P_{n}(x) = \sum_{i=0}^{n} c_{i}^{2} x^{2i}$$

and

$$r(x, y) := \frac{\mathbb{E}P_{n}(x)P_{n}(y)}{\sqrt{V(x)} \sqrt{V(y)}} = \frac{\sum_{i=0}^{n} c_{i}^{2} x^{i} y^{i}}{\sqrt{\left(\sum_{i=0}^{n} c_{i}^{2} x^{2i}\right) \left(\sum_{i=0}^{n} c_{i}^{2} y^{2i}\right)}}.$$  

We will use the following lemma that bounds the probability that the polynomial $P_{n}$ has many roots in a small interval.

**Lemma 5.1.** Assume that the random variables $\xi_{i}$ are iid standard Gaussian. There exists a constant $C_{0}$ such that for any $0 < s < 1$, any $k, l \geq 2$, $1 - \frac{1}{C_{0}} \leq x < t < 1$ and $y, z \in (x, t)$ satisfying

$$\log \frac{1 - x}{1 - z} = \log \frac{1 - y}{1 - t} = \delta$$

for some $\delta \in (0, 1/2C_{0})$, we have

$$\mathbb{P}(N_{n}(x, y) \geq k) \ll (C_{0})^{ks}$$

and

$$\mathbb{P}(N_{n}(x, y) \geq k, N_{n}(z, t) \geq l) \ll (C_{0})^{2ks} + (C_{0})^{2ls} + \frac{(C_{0})^{(k+l)s}}{\sqrt{1 - r^{2}(y, t)}}.$$  

where the implicit constants depend only on $s$, not on $k, l, x, y, \delta$.

We now prove (39), assuming Lemma 5.1. In fact, we will use only (45); (46) is needed in later sections.

**Proof of (39).** By (45), for every interval $[x, y]$ with

$$1 - b_{n} \leq x < y < 1 - \frac{a_{n}}{n}$$

and $\log \frac{1 - x}{1 - y} = \frac{1}{C}$,

where $C$ is a sufficiently large constant, we have

$$\mathbb{E}N_{n}^{k}(x, y) \leq \mathbb{E}N_{n}(x, y) + \sum_{j=2}^{\infty} j^{k}\mathbb{P}(N_{n}(x, y) = j) \ll \mathbb{E}N_{n}(x, y) + \sum_{j=2}^{\infty} j^{k} 2^{-j} = \mathbb{E}N_{n}(x, y) + O(1).$$

Dividing the interval $(1 - b_{n}, 1 - \frac{a_{n}}{n})$ into $O\left(\log \frac{n}{a_{n}} + \log b_{n}\right) = O\left(\log \frac{nb_{n}}{a_{n}}\right)$ intervals that satisfy (47), we obtain

$$\mathbb{E}N_{n}^{k} \left(1 - b_{n}, 1 - \frac{a_{n}}{n}\right) \ll \left(\log \frac{nb_{n}}{a_{n}}\right)^{k-1} \mathbb{E}N_{n} \left(1 - b_{n}, 1 - \frac{a_{n}}{n}\right) + \left(\log \frac{nb_{n}}{a_{n}}\right)^{k-1}.$$  

So, (39) follows from (48) and the following

$$\mathbb{E}N_{n} \left(1 - b_{n}, 1 - \frac{a_{n}}{n}\right) \ll \max\{1, \log(nb_{n})\}.$$
which can be deduced from
\begin{equation}
\mathbb{E}N_n \left( 1 - b_n, 1 - \frac{C}{n} \right) \ll \log(nb_n) \quad \text{if} \quad b_n \geq C/n
\end{equation}
and
\begin{equation}
\mathbb{E}N_n \left( 1 - \frac{C}{n}, 1 - \frac{a_n}{n} \right) \ll 1.
\end{equation}

To prove (50), let \( c_{i,\rho} := \sqrt{\frac{(2\rho+1)\cdots(2\rho+i)}{i!}} \). We have \( c_{i,\rho} = \Theta(c_i) \) for all \( i \geq N_0 \) thanks to assumption A3.

Using Kac-Rice formula (20), we have
\begin{equation}
\mathbb{E}N_n \left( 1 - b_n, 1 - \frac{C}{n} \right) \ll \int_{1-b_n}^{1-C/n} \sqrt{\sum_{i=0}^{n} \sum_{j=i+1}^{n} c_{i,\rho}^2 c_{j,\rho}^2 (j-i)^2 t^{2j+2i-2} \over \sum_{i=0}^{n} c_{i,\rho}^2 t^{2i}} \, dt.
\end{equation}

We use [6, Lemma 10.3] with \( h(k) = c_{i,\rho}^2 \) which estimates the above integrand uniformly over the interval \((1 - b_n, 1 - \frac{C}{n})\) and asserts that
\begin{equation}
\frac{\sqrt{\sum_{i=0}^{n} \sum_{j=i+1}^{n} c_{i,\rho}^2 c_{j,\rho}^2 (j-i)^2 t^{2j+2i-2} \over \sum_{i=0}^{n} c_{i,\rho}^2 t^{2i}}}{2n} \ll \frac{2^{\rho+1}}{2\pi(1-t)} + (1-t)^{\rho-1/2} + \frac{1}{n(1-t)^2},
\end{equation}
which is \( \ll \frac{1}{1-t} \) by the assumption \( \rho > -1/2 \).

That gives (50) because
\begin{equation}
\mathbb{E}N_n \left( 1 - b_n, 1 - \frac{C}{n} \right) \ll \int_{1-b_n}^{1-C/n} \frac{1}{1-t} \, dt \ll \log n + \log b_n = \log(nb_n).
\end{equation}

For (51), we use the same bound as in (41) to obtain
\begin{equation}
\mathbb{E}N_n \left( 1 - \frac{C}{n}, 1 - \frac{a_n}{n} \right) \ll \int_{1-C/n}^{1-a_n/n} n \, dt \ll 1.
\end{equation}
This proves (51) and completes the proof of (39).

Proof of Lemma 5.1. We start by proving (46). By Rolle’s theorem and the fundamental theorem of calculus, if \( P \) has at least \( k \) zeros in the interval \((x, y)\) then
\begin{equation}
|P_n(y)| \leq \int_x^y \int_x^{y_1} \ldots \int_x^{y_{k-1}} |P_n^{(k)}(y_k)| \, dy_k \ldots \, dy_1 =: I_{x,y}.
\end{equation}

Therefore,
\begin{align*}
\mathbb{P}(N_n(x,y) \geq k, N_n(z,t) \geq l) \leq & \ \mathbb{P} \left( I_{x,y} \geq \varepsilon_1 \sqrt{V(y)} \right) + \mathbb{P} \left( I_{z,t} \geq \varepsilon_2 \sqrt{V(t)} \right) \\
& + \ \mathbb{P} \left( |P_n(y)| \leq \varepsilon_1 \sqrt{V(y)}, |P_n(t)| \leq \varepsilon_2 \sqrt{V(t)} \right)
\end{align*}

where \( \varepsilon_1 := (C_0\delta)^{k+}, \varepsilon_2 := (C_0\delta)^{l+} \) and
\begin{equation}
V(x) = \text{Var} \ P_n(x) = \sum_{i=0}^{n} c_i^2 x^{2i}.
\end{equation}
By (A3), we have the following estimate whose proof is deferred to Appendix 8.4 as it is merely algebraic:

$$(56) \quad V(x) = \frac{\Theta(1)}{(1 - x + 1/n)^{2\rho+1}} \quad \forall x \in (1 - 1/C, 1).$$

Since $\left(\frac{P_n(y)}{\sqrt{V(y)}}, \frac{P_n(t)}{\sqrt{V(t)}}\right)$ is a Gaussian vector with mean 0 and covariance matrix $\begin{bmatrix} 1 & r(y, t) \\ r(y, t) & 1 \end{bmatrix}$, we have

$$\mathbb{P}\left( |P_n(y)| \leq \varepsilon_1 \sqrt{V(y)}, |P_n(t)| \leq \varepsilon_2 \sqrt{V(t)} \right) \ll \frac{\varepsilon_1 \varepsilon_2}{\sqrt{1 - r^2(y, t)}}.$$ 

It remains to show that

$$(57) \quad \mathbb{P}\left( I_{x,y} \geq \varepsilon_1 \sqrt{V(y)} \right) \ll (C_0 \delta)^{2k \rho}.$$ 

Since $0 < s < 1$, there exists $h > 0$ such that $s = \frac{2 + h}{4 \cdot h}$. By Markov’s inequality, we have

$$\mathbb{P}\left( I_{x,y} \geq \varepsilon_1 \sqrt{V(y)} \right) \leq \mathbb{E} \left( \int_x^y \int_x^{y_1} \cdots \int_x^{y_{k-1}} |P_n^{(k)}(y_k)|^2 dy_k \right)^{2 \rho}.$$ 

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

$$\left( \frac{(y - x)^k}{k!} \right)^{1+h} \mathbb{E} \int_x^y \int_x^{y_1} \cdots \int_x^{y_{k-1}} |P_n^{(k)}(y_k)|^2 dy_k \cdots dy_{k-1}$$

and so

$$(58) \quad \mathbb{P}\left( I_{x,y} \geq \varepsilon_1 \sqrt{V(y)} \right) \leq \left( \frac{(y - x)^k}{k!} \right)^{2 \rho} \sup_{w \in (x,y)} \mathbb{E} |P_n^{(k)}(w)|^{2 \rho}.$$ 

For each $w \in (x,y)$, since $P_n^{(k)}(w)$ is a Gaussian random variable, using the hypercontractivity inequality for the Gaussian distribution (see, for example, [4, Corollary 5.21]), we have for some constant $C$,

$$\mathbb{E} |P_n^{(k)}(w)|^{2 \rho} \ll \left( \mathbb{E} |P_n^{(k)}(w)|^{2 \rho} \right)^{\frac{2 \rho}{2 \rho + 2 k + 1}} \ll \left( \frac{C^k(k!)^2}{(1 - y + 1/n)^{2 \rho + 2 k + 1}} \right)^{\frac{2 \rho}{2 \rho + 2 k + 1}}$$

where in the last inequality, we used an estimate similar to (56).

Plugging this and (56) into (58), we obtain

$$\mathbb{P}\left( I_{x,y} \geq \varepsilon_1 \sqrt{V(y)} \right) \ll \left( \frac{1 - y + 1/n}{2 \rho + 2 k + 1} \right)^{\frac{2 \rho}{2 \rho + 2 k + 1}} \left( \frac{C^k(k!)^2}{(1 - y + 1/n)^{2 \rho + 2 k + 1}} \right)^{\frac{2 \rho}{2 \rho + 2 k + 1}}$$

which gives

$$\mathbb{P}\left( I_{x,y} \geq \varepsilon_1 \sqrt{V(y)} \right) \ll \frac{1}{\varepsilon_1^{2 \rho}} \left( C \frac{y - x}{1 - y + 1/n} \right)^{k(2 \rho)} \ll \frac{1}{\varepsilon_1^{2 \rho}} \left( C \frac{y - x}{1 - y} \right)^{k(2 \rho)}.$$ 

Using $\varepsilon_1 = (C_0 \delta)^{ks}$, $\frac{y - x}{1 - y} = \frac{1 - x}{1 - y} - 1 = e^s - 1 \leq 2\delta$ for $\delta \leq \frac{1}{2C_0}$ and $s = \frac{2 \rho}{4 \cdot h}$, we get

$$\mathbb{P}\left( I_{x,y} \geq \varepsilon_1 \sqrt{V(y)} \right) \ll \frac{1}{(C_0 \delta)^{k s (2 \rho)}} (2 \rho C \delta)^{k(2 \rho)} \ll (C_0 \delta)^{2 k s}$$

by choosing $C_0 \geq 2C$. This proves (46).
The inequality (45) is obtained by the same reasoning:
\[
P(N_n(x,y) \geq k) \leq P(I_{x,y} \geq \varepsilon_1 \sqrt{V(y)}) + P(|P_n(y)| \leq \varepsilon_1 \sqrt{V(y)}) \ll (C_0 \delta)^{2ks} + \varepsilon_1 \ll (C_0 \delta)^{ks}.
\]
This completes the proof of Lemma 5.1. □

6. Proof of Lemma 2.4

Since the lemma only involves Gaussian random variables \(\tilde{\xi}_i\), we simplify the notation and write \(\xi_i\) for \(\tilde{\xi}_i\) and \(N_n(S)\) for \(\tilde{N}_n(S)\) (this helps us to avoid multiple superscripts later on). Thus, for this section, \(\xi_i \sim N(0,1)\) for all \(i\).

We will adapt the argument in Maslova [25], which is to approximate the number of roots by a sum of independent random variables. Since the random variables \(\xi_i\) are now standard Gaussian, numerous technical steps in [25], which may be impossible to reproduce without having \(c_0 = \cdots = c_n = 1\), can be greatly simplified and applied to our general setting thanks to special properties of Gaussian variables.

**Step 1: Approximate the number of real roots by the number of sign changes.**

Let \(V\) and \(r\) be defined as in (43) and (44). Lemma 5.1 asserts that in a small interval, it is unlikely that the polynomial \(P_n\) has more than 1 root. If \(P_n\) has at most 1 root in an interval \((a,b)\) and does not vanish at \(a\) and \(b\) then \(N_n(a,b) = 1\) if \(P_n(a)\) and \(P_n(b)\) have different signs and \(N_n(a,b) = 0\) otherwise. Hence, on a small interval \((a,b)\), it is reasonable to approximate \(N_n(a,b)\) by the number of sign changes:

\[
N_{n}^{\text{sign}}(a,b) = \frac{1}{2} - \frac{1}{2} \text{sign}(P_n(a)P_n(b))
\]

where
\[
\text{sign}(x) := \begin{cases} 
1 & \text{if } x > 0, \\
0 & \text{if } x = 0, \\
-1 & \text{if } x > 0.
\end{cases}
\]

The following lemma estimates the accuracy of this approximation for a long interval.

**Lemma 6.1 (Approximate by sign changes).** Assume that the \(\xi_i\) are iid standard Gaussian. For any positive constant \(\varepsilon\), there exist constants \(C, C'\) such that the following holds. Let \(T > 1/C\) and \(a, b\) be such that \(1 - a_n \leq a < b \leq 1 - b_n\) and \(\log \frac{1-a}{1-b} = T\). Let \(j_0 = \delta^{-1} \log(1-a)^{-1}\) and \(j_1 = \delta^{-1} \log(1-b)^{-1}\) where \(\delta\) is any number with
\[
\exp(-((\log \log n)^{1+\varepsilon})) < \delta < 1/C.
\]

Assume (without loss of generality) that \(j_0\) and \(j_1\) are integers and let \(x_j = 1 - \exp(-j\delta)\) for all \(j = j_0, \ldots, j_1\). Let
\[
S = S_{a,b,\delta} = N_n(a,b) = \sum_{j=j_0}^{j_1} N_n[x_j, x_{j+1}] \quad \text{and} \quad S_{a,b,\delta}^{\text{sign}} = S_{a,b,\delta}^{\text{sign}} = \sum_{j=j_0}^{j_1} N_n^{\text{sign}}(x_j, x_{j+1}).
\]

Then
\[
E(S - S_{a,b,\delta}^{\text{sign}})^2 \leq C' T^2 \delta^{1-\varepsilon}.
\]
Step 2: Truncate the polynomial $P_n$ to get independence.

Assume Lemma 6.1 for a moment. The next trick (following Maslova) is to show that $N^\text{sign}(x,y)$ and $N^\text{sign}(z,t)$ (in some rough sense) are independent, whenever the intervals $(x,y)$ and $(z,t)$ are relatively far apart. This allows us to approximate $N_n(J)$ by a sum of independent random variables, from which we can derive a Central Limit Theorem.

For any $x \in [1-a_n, 1-b_n]$, let

$$A_x = \log(1-x)^{-1}, \quad m_x = (1-x)^{-1}A_x^{-\alpha}, \quad \text{and} \quad M_x = \alpha(1-x)^{-1}\log A_x$$

where $\alpha$ is a large constant to be chosen.

Define a truncated version of $P_n$ by

$$Q(x) = \sum_{j=m_x}^{M_x} c_j \xi_j x^j.$$ 

We get $Q$ from $P$ by a truncation in which the truncation points $m_x$ and $M_x$ depend on the value of $x$. Let

$$\rho' = \min\{1, 1 + 2\rho\} > 0.$$ 

The following lemma asserts that $Q$ is a good approximation of $P_n$ and that $Q(x)$ and $Q(y)$ are independent when $x$ and $y$ are far apart. We defer the routine proof of this lemma to Appendix S.6.

**Lemma 6.2.** For every $x \in [1-a_n, 1-b_n]$, it holds that

$$0 \leq \text{Var} \; P_n(x) - \text{Var} \; Q(x) = \text{Var} \; (P_n(x) - Q(x)) \ll A_x^{-\alpha \rho'} \mathbb{E} P_n^2(x),$$

Moreover, if $1-a_n \leq x < y \leq 1-b_n$ and if $\log \frac{1-x}{1-y} \geq 2\alpha \log \log n$ then $Q(x)$ and $Q(y)$ are independent because $M_x < m_y$.

Let

$$N^\text{trun}_{P_n}(x,y) = N^\text{trun}(x,y) := \frac{1}{2} - \frac{1}{2} \text{sign}(Q(x)Q(y))$$

be the sign change of $Q$ on the interval $(x,y)$. In the next lemma, we show that $N^\text{trun}$ is a good approximation of the corresponding sign change $N^\text{sign}_n$ of $P_n$ defined in (60).

**Lemma 6.3** (Approximation by truncation I). Assume that the $\xi_i$ are iid standard Gaussian. Let $C$ be any positive constant. Let $1-a_n \leq x < y \leq 1-b_n$ with $\log \frac{1-x}{1-y} \leq 1/C$. Then

$$\mathbb{E} \left( N^\text{sign}_n(x,y) - N^\text{trun}(x,y) \right)^2 \ll A_x^{-\alpha \rho'/3}.$$ 

Assuming Lemma 6.3 for a moment, we proceed to obtain

**Lemma 6.4** (Approximation by truncation II). Assume that the $\xi_i$ are iid standard Gaussian. There exist constants $C,C'$ such that the following holds. Let $T > 1/C$ and $a,b$ be such that $1-a_n \leq a < b \leq 1-b_n$ and $\log \frac{1-a}{1-b} = T$. Let $j_0 = \delta^{-1} \log(1-a)^{-1}$ and $j_1 = \delta^{-1} \log(1-b)^{-1}$
where \( \delta \) is any number in (0, 1/C). Assume (without loss of generality) that \( j_0 \) and \( j_1 \) are integers and let \( x_j = 1 - \exp(-j\delta) \) for all \( j = j_0, \ldots, j_1 \). Let

\[
S_{\text{sign}}^{\text{sign}} = S_{a,b,\delta}^{\text{sign}} = \sum_{j=j_0}^{j_1} N_n^{\text{sign}}(x_j, x_{j+1}) \quad \text{and} \quad S_{\text{trun}}^{\text{trun}} = S_{a,b,\delta}^{\text{trun}} = \sum_{j=j_0}^{j_1} N_n^{\text{trun}}(x_j, x_{j+1}).
\]

Then

\[
\mathbb{E}(S_{\text{trun}}^{\text{trun}} - S_{\text{sign}}^{\text{sign}})^2 \leq C'\delta^{-2}T^2 \left( \log \frac{1}{a_n} \right)^{-\alpha \rho' / 3} = C'\delta^{-2} \left( \log \frac{1-a}{1-b} \right)^2 \left( \log \frac{1}{a_n} \right)^{-\alpha \rho' / 3}.
\]

Proof. By Lemma 6.3, we have

\[
\mathbb{E} \left( S_{\text{trun}}^{\text{trun}} - S_{\text{sign}}^{\text{sign}} \right)^2 \leq \left( \sum_{j=j_0}^{j_1} \left[ \mathbb{E}(N_n^{\text{trun}}(x_j, x_{j+1}) - N_n^{\text{sign}}(x_j, x_{j+1}))^2 \right]^{1/2} \right)^2 \leq \left( \sum_{j=j_0}^{j_1} \frac{1}{a_n} \right)^{-\alpha \rho' / 6} \leq \left( \sum_{j=j_0}^{j_1} (j\delta)^{-\alpha \rho' / 6} \right)^2 \leq \left( \sum_{j=j_0}^{j_1} \frac{1}{a_n} \right)^{-\alpha \rho' / 6} \leq T \left( \log \frac{1}{a_n} \right)^{-\alpha \rho' / 6}.
\]

By the definition of \( j_0 \) and \( j_1 \), we get

\[
\sum_{j=j_0}^{j_1} (j\delta)^{-\alpha \rho' / 6} \leq \delta^{-\alpha \rho' / 6} (j_1 - j_0) \leq \delta^{-1 - \alpha \rho' / 6} T j_0^{-\alpha \rho' / 6} \leq \delta^{-1 - \alpha \rho' / 6} T \left( \log \frac{1}{a_n} \right)^{-\alpha \rho' / 6},
\]

proving Lemma 6.4. \( \square \)

The following lemma controls the forth moment of \( S_{\text{trun}}^{\text{trun}} \).

**Lemma 6.5 (Bounded forth moment).** Under the setting of Lemma 6.4 and an additional assumption that \( \delta \geq \left( \log \frac{1}{a_n} \right)^{-\alpha \rho' / 24} \), we have

\[
(64) \quad \mathbb{E} \left( S_{a,b,\delta}^{\text{trun}} - \mathbb{E} S_{a,b,\delta}^{\text{trun}} \right)^4 \ll T^2 (\log \log n)^2 = \left( \log \frac{1-a}{1-b} \right)^2 (\log \log n)^2.
\]

Assuming Lemma 6.5, we are now ready for the proof of Lemma 2.4.

**Step 3: Proof of Lemma 2.4.** Using the previous two steps, we shall approximate \( N_n(3) \) by a sum of independent random variables to prove that it satisfies the CLT. We again recall that in this proof, the \( \xi_i \) are iid standard Gaussian as mentioned at the beginning of this section. Recall the hypothesis (6) that

\[
(65) \quad (\log n)^2 / n \leq b_n \leq a_n \leq \exp\left( - (\log n)^c \right), \quad \log \frac{a_n}{b_n} = \Theta(\log n), \quad \text{and} \quad \text{Var} \ N_n(3) = \Omega(\log n).
\]

In particular, \( a_n \) satisfies Condition (2). Let \( \alpha, \beta \) be any constants satisfying

\[
(66) \quad \beta \geq 3 \quad \text{and} \quad 2\beta + 3 \leq \alpha \rho' / 24.
\]

Let

\[
(67) \quad T := \log \frac{a_n}{b_n} = \Theta(\log n), \quad \delta := (\log n)^{-\beta}, \quad j_0 := \delta^{-1} \log \frac{1}{a_n} \quad \text{and} \quad j_1 := \delta^{-1} \log \frac{1}{b_n}.
\]
We have \( j_1 - j_0 = \delta^{-1}T \). Let
\[
q := \delta^{-1}T^{1/8} \quad \text{and} \quad p := \delta^{-1}T^{1/2}.
\]
Observe that \( q = o(p) \) and \( q \) grows with \( n \). For simplicity, we will assume that \( j_0, j_1, p \) and \( q \) are integers. In the case that they are not, we only need to replace them by their integer part. As before, let \( x_j = 1 - \exp(-j\delta) \) for \( j = j_0, \ldots, j_1 \).

Let \( N_{P_n}^{\text{trun}}(x_j, x_{j+1}) \) be defined as in Eq. (68). By Lemmas 6.1 and 6.4, we can approximate \( N_n(\mathcal{I} \cap (0, 1)) \) by
\[
S_1^{\text{trun}} := S_1^{\text{trun}, 1} + \sum_{j=j_0}^{j_1-1} N_{P_n}^{\text{trun}}(x_j, x_{j+1})
\]
and get an error term
\[
\mathbb{E} \left( N_n(\mathcal{I} \cap (0, 1)) - S_1^{\text{trun}} \right)^2 \ll T^2\delta^{1-\varepsilon} + T^2\delta^{-2} \left( \log \frac{1}{\delta n} \right)^{-\alpha/3} = o(\log n)
\]
where in the last inequality, we used (65) and (66).

Combining this with the assumption that \( \text{Var} \ N_n(\mathcal{I}) = \Omega(\log n) \), we get
\[
\mathbb{E} \left( N_n(\mathcal{I} \cap (0, 1)) - S_1^{\text{trun}} \right)^2 = o(\log n) = o(\text{Var} \ N_n(\mathcal{I})).
\]

Similarly, for the interval \( \mathcal{I} \cap (-1, 0) \), we approximate the number of real roots by
\[
S_2^{\text{trun}} := \sum_{j=j_0}^{j_1-1} N_{P_n}^{\text{trun}}(-x_{j+1}, -x_j).
\]

And for the intervals \( \mathcal{I} \cap (1, \infty) \) and \( \mathcal{I} \cap (-\infty, -1) \), we respectively use
\[
S_3^{\text{trun}} := \sum_{j=j_0}^{j_1-1} N_{R_n}^{\text{trun}}(x_j, x_{j+1}) \quad \text{and} \quad S_4^{\text{trun}} := \sum_{j=j_0}^{j_1-1} N_{R_n}^{\text{trun}}(-x_{j+1}, -x_j)
\]
where \( R_n(x) = \frac{e_n}{c_n} P_n(x^{1-1}) = \sum_{i=0}^{n} \frac{c_{n-i}}{c_n} \xi_{n-i} x^i \). Let \( S^{\text{trun}} := \sum_{k=1}^{4} S_k^{\text{trun}} \). We note that all of the lemmas proven earlier in this section hold for \( R_n \) in place of \( P_n \) (with the value of \( \rho \) being changed to 0 as in Remark 3.8). From (68) and its analog for \( S_2^{\text{trun}}, S_3^{\text{trun}}, S_4^{\text{trun}} \), we have
\[
\mathbb{E}(N_n(\mathcal{I}) - S^{\text{trun}})^2 = o(\log n) = o(\text{Var} \ N_n(\mathcal{I})).
\]

Making use of Lemma 6.2, we now approximate \( S^{\text{trun}} \) by a sum of independent random variables \( Z_k, W_k \) as follows. Let
\[
Z_k = \sum_{j=j_0+(k+1)p+q}^{j_0+(k+1)p+kq-1} \left( N_{P_n}^{\text{trun}}(x_j, x_{j+1}) + N_{P_n}^{\text{trun}}(-x_{j+1}, -x_j) \right),
\]
and
\[
W_k = \sum_{j=j_0+(k+1)p+q}^{j_0+(k+1)p+kq-1} \left( N_{R_n}^{\text{trun}}(x_j, x_{j+1}) + N_{R_n}^{\text{trun}}(-x_{j+1}, -x_j) \right), \quad k = 0, \ldots, l - 1
\]
where
\[
l = \frac{j_1 - j_0}{p + q} = \Theta(T^{1/2}).
\]
By Lemma 6.2 the random variables $Z_0, \ldots, Z_{l-1}$ are mutually independent because $q\delta = T^{1/8} \geq 2\alpha \log \log n$. Similarly for the random variables $W_0, \ldots, W_{l-1}$. Moreover, all random variables $Z_0, \ldots, Z_{l-1}, W_0, \ldots, W_{l-1}$ are mutually independent because the $Z_s$ only involve the random variables $\xi_r$ where $r \leq M_1-b_0 \leq n/2$ (by the definition of $Z_s$ and the left-most inequality in (65)) while the $W_s$ only involve the random variables $\xi_{n-r}$ where, again, $r \leq M_1-b_0 \leq n/2$.

To evaluate the accuracy of the approximation of $S_{\text{trun}}$ by $\sum_k (Z_k + W_k)$, consider

$$S_{\text{trun}} - \sum_{k=0}^{l-1} (Z_k + W_k) = \sum_{k=0}^{l-1} (X_k + Y_k)$$

where

$$X_k = \sum_{j=j_0+(k+1)p+kq-1}^{j_0+(k+1)p+(k+1)q-1} (N_{\text{trun}}^{trun}(x_j, x_{j+1}) + N_{\text{trun}}^{trun}(-x_{j+1}, -x_j)),$$

for $k = 0, 1, \ldots, l-1$, and $Y_k$ are defined similarly with respect to $R_n$.

By Lemma 6.2, the random variables $X_0, \ldots, X_{l-1}, Y_0, \ldots, Y_{l-1}$ are also mutually independent. Note that each $X_k, Y_k$ is of the form $S_{a,b,\delta}^{\text{trun}}$ defined in Lemma 6.4 for some $a$ and $b$ with $\log \frac{1-b}{1-a} = q\delta = T^{1/8}$. By (66) and the definition of $\delta$ in (67), $\delta = (\log n)^{-\beta} \geq \left(\log \frac{1}{a_n}\right)^{-\alpha/24}$; this allows us to use Lemma 6.5 to get

$$E(X_k - EX_k)^4 \ll q^2\delta^2(\log \log n)^2 \quad \text{for all } k = 0, \ldots, l-1.$$

One can obtain a similar estimate for $Y_k$. Thus, the error term of the approximation of $S_{\text{trun}}$ by $\sum_{k=0}^{l-1} (Z_k + W_k)$ has variance

$$\text{Var} \left( \sum_{k=0}^{l-1} (X_k + Y_k) \right) = \sum_{k=0}^{l-1} \text{Var} X_k + \sum_{k=0}^{l-1} \text{Var} Y_k \ll \sum_{k=0}^{l-1} q\delta \log \log n = o(\log n).$$

Combining this with (68), we get

$$\text{Var} \left( N_n(\overline{J}) - \sum_{k=0}^{l-1} (Z_k + W_k) \right) = o(\log n) = o(\text{Var} \ N_n(\overline{J})). \quad (69)$$

The sum $\sum_{k=0}^{l-1} (Z_k + W_k)$ is a sum of independent random variables satisfying forth moment bound

$$\sum_{k=0}^{l-1} E(Z_k - EZ_k)^4 + \sum_{k=0}^{l-1} E(W_k - EW_k)^4 \ll \sum_{k=0}^{l-1} p^2\delta^2(\log \log n)^2 = o(\log^2 n) = o \left( \text{Var} \sum_{k=0}^{l-1} (Z_k + W_k) \right)^2$$

where in the first inequality, we used Lemma 6.5. By the Lyapunov Central Limit Theorem (see for example, [28]), the sum $\sum_{k=0}^{l-1} (Z_k + W_k)$ satisfies the Central Limit Theorem.

This and (69) imply that $N_n(\overline{J})$ also satisfies the Central Limit Theorem, completing the proof of Lemma 2.4.

Proof of Lemma 6.7. Note that $x_{j_0} = a, x_{j_1} = b$, and $j_1 - j_0 = \delta^{-1}T$. 

\hfill \square
We have

\[ \mathbb{E}(S - S^{\text{sign}})^2 = \sum_{i=j_0}^{j_1-1} \mathbb{E}(N_i - N_i^{\text{sign}})^2 + 2 \sum_{j_0 < i < j_1} \mathbb{E}(N_i - N_i^{\text{sign}})(N_j - N_j^{\text{sign}}) \]

where \( N_j := N_n(x_j, x_{j+1}) \) and \( N_j^{\text{sign}} := N_n^{\text{sign}}(x_j, x_{j+1}) \).

By Lemma 5.1, we have

\[ \sum_{i=j_0}^{j_1-1} \mathbb{E}(N_i - N_i^{\text{sign}})^2 \leq \sum_{i=j_0}^{j_1-1} \sum_{k=2}^{n} k^2 \mathbb{P}(N_i = k) \ll T \delta^{-1} \sum_{k=2}^{n} k^2 (C_0 \delta^{k(1-\varepsilon)/2}) \ll T \delta^{1-\varepsilon}. \]

For each \( j_0 \leq i < j \leq j_1 - 1 \), we have

\[ \mathbb{E}(N_i - N_i^{\text{sign}})(N_j - N_j^{\text{sign}}) \leq \sum_{k,l=2}^{n} kl \mathbb{P}(N_i = k, N_j = l). \]

Let \( k_0 := \delta^{-1/100} \). We split the right-hand side into three sums: \( 2 \leq k, l \leq k_0 \) for the first sum, \( k_0 < k \leq n \) and \( 2 \leq l \leq n \) for the second sum, and \( 2 \leq k \leq n \) and \( 0 < l \leq n \) for the third sum, and denote the corresponding sums by \( K_1, K_2, K_3 \), respectively.

By Lemma 5.1 letting \( r_{ij} := r(x_{i+1}, x_{j+1}) \) gives

\[ K_1 \ll k_0^2 \left[ (C_0 \delta)^{4(1-\varepsilon)} + (C_0 \delta)^{4(1-\varepsilon)} \right] \ll \frac{\delta^3}{\sqrt{1 - r_{ij}^2}}. \]

For \( K_2 \), we use Hölder’s inequality to get

\[ K_2 \leq \mathbb{E}(N_i N_j \mathbb{1}_{N_i \geq k_0+1} \mathbb{1}_{N_j \geq 2}) \leq \left( \mathbb{E}N_i^2 \mathbb{1}_{N_i \geq k_0+1} \right)^{1/2} \left( \mathbb{E}N_j^2 \mathbb{1}_{N_j \geq 2} \right)^{1/2} \leq k_0^{-1/2} \left( \mathbb{E}N_i^2 \mathbb{1}_{N_i \geq 2} \right)^{1/2} \ll k_0^{-h+1} \delta^2 \ll \delta^3 \]

where \( h \) is a sufficiently large constant and in the next to last inequality, we used Lemma 5.1 in a similar way as in (70). Similarly, \( K_3 \ll \delta^3 \). Hence,

\[ \mathbb{E}(N_i - N_i^{\text{sign}})(N_j - N_j^{\text{sign}}) \ll \delta^3 \frac{\delta^3}{\sqrt{1 - r_{ij}^2}}, \]

and so

\[ \mathbb{E}(S - S^{\text{sign}})^2 \ll T \delta^{1-\varepsilon} + \sum_{j_0 < i < j_1} \left( \delta^3 \frac{\delta^3}{\sqrt{1 - r_{ij}^2}} \right) \ll T \delta^{1-\varepsilon} + \delta^3 \sum_{j_0 < i < j_1} 1. \]

To complete the proof of the lemma, it remains to bound \( 1 - r_{ij}^2 \) from below. For each \( 0 \leq k \leq n \), let

\[ c_{k, \rho} = \sqrt{(k+2\rho) \cdots (1+2\rho) / k!}. \]
By Condition (A3), \( c_k = \Theta(c_{k,\rho}) \) for all \( k \geq N_0 \) and thus, for all \( x, y \in [a, b], \)

\[
V(x) = \sum_{k=0}^{n} c_k^2 x^{2k} = \Theta \left( \sum_{k=0}^{n} c_{k,\rho}^2 x^{2k} \right)
\]

and

\[
1 - r^2(x, y) = \sum_{0 \leq i < k \leq n} c_i^2 c_{k-i}^2 (xy - x^i y^k - x^k y^i)^2 \Theta \left( \sum_{0 \leq i < k \leq n} c_{i,\rho}^2 c_{k-i,\rho}^2 (xy - x^i y^k - x^k y^i)^2 \right).
\]

Therefore, in order to bound \( 1 - r^2 \) from below, it suffices to assume that \( c_k = c_{k,\rho} \) for all \( 0 \leq k \leq n \) for the rest of the proof of Lemma 6.1.

For \( c_k = c_{k,\rho} \), we have for every \( x \in [1 - a_n, 1 - b_n], \)

\[
V(x) = \frac{1 + O(\varepsilon_0)}{(1 - x^2)^2 + 1}.
\]

where \( \varepsilon_0 = \exp \left( -(\log \log n)^{1+2e} \right) \). We defer the simple verification of (73) and (74) to Appendix 8.5.

Letting \( x = x_{i+1} \) and \( y = x_{j+1} \) yields

\[
r_{ij} = \frac{V(\sqrt{xy})}{\sqrt{V(x)V(y)}} = (1 + O(\varepsilon_0)) \frac{\sqrt{(1 - x^2)(1 - y^2)}}{(1 - xy)}^{2\rho + 1}.
\]

Let \( s_{ij} := \sqrt{(1 - x^2)(1 - y^2)} \). To estimate \( 1 - r^2 \), let us first estimate \( 1 - s^2 \). We have

\[
1 - s^2 = \frac{(x - y)^2}{(1 - x + x)(1 - y)^2} \geq \frac{(e^{(j-i)\delta} - 1)^2}{(e^{(j-i)\delta} + 1)^2} \geq \frac{(j - i)^2 \delta^2}{(e^{(j-i)\delta} + 1)^2}.
\]

Thus, if \( (j - i)\delta \leq 1 \) then \( 1 - s^2 \gg (j - i)^2 \delta^2 \) and if \( (j - i)\delta \geq 1 \) then \( 1 - s^2 = \frac{(e^{(j-i)\delta} - 1)^2}{(e^{(j-i)\delta} + 1)^2} \gg 1 \).

Combining this with the assumption that \( \delta \geq \exp \left( -(\log \log n)^{1+e} \right) \), we have \( \varepsilon_0 = o \left( 1 - s^2 \right) \) for all \( i < j \). This implies

\[
1 - r^2 = 1 - s^2(2\rho + 1) + o(1 - s^2) = \Theta(1 - s^2) = \Theta \left( \frac{(x - y)^2}{(1 - xy)^2} \right)
\]

and

\[
\sum_{j_0 \leq i < j \leq j_1 - 1} \frac{1}{\sqrt{1 - r^2}} \ll \sum_{\substack{i < j < i+\delta - 1}} \frac{1}{\sqrt{1 - s^2}} + \sum_{j \geq i+\delta - 1} \frac{1}{\sqrt{1 - s^2}} \ll \sum_{\substack{i < j < i+\delta - 1}} \frac{1}{(j - i)\delta} + \sum_{j \geq i+\delta - 1} 1 \ll T\delta^{-2} \log \delta^{-1} + T^2 \delta^{-2}.
\]

Plugging this into (72), we obtain

\[
\mathbb{E}(S - S^{\text{sign}})^2 \ll T\delta^{-1} + T^2 \delta + T\delta \log \delta^{-1} \ll T^2 \delta^{1-\varepsilon},
\]

completing the proof of Lemma 6.1. \( \square \)
Proof of Lemma 6.3. Using the formula
\[ \text{sign}(a) = \frac{1}{\pi} \int_{\mathbb{R}} t^{-1} \sin(ta) dt, \]
we have
\[ N_n^{\text{sign}}(x,y) - N_n^{\text{trun}}(x,y) = \frac{1}{2\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} t^{-1} u^{-1} \left( \sin(t\bar{Q}(x)) \sin(u\bar{Q}(y)) - \sin(t\bar{P}_n(x)) \sin(u\bar{P}_n(y)) \right) dt du \]
where
\[ \bar{Q}(x) := \frac{Q(x)}{\sqrt{V(x)}}, \quad \bar{Q}(y) := \frac{Q(y)}{\sqrt{V(y)}}, \quad \bar{P}_n(x) := \frac{P_n(x)}{\sqrt{V(x)}}, \quad \text{and} \quad \bar{P}_n(y) := \frac{P_n(y)}{\sqrt{V(y)}}. \]

Decompose the plane \( \mathbb{R} \times \mathbb{R} \) of \((t,u)\) into two regions: the square \( \{ (t,u) : A_x^{-\alpha \rho/6} \leq |t|, |u| \leq A_x^{-\alpha \rho/3} \} \) and its complement. We denote the corresponding integrals on these regions by \( I_1 \) and \( I_2 \), respectively.

First, we show that the contribution from \( I_2 \) is negligible. Indeed, using the estimates
\[ \left| \int_{|t| \leq \varepsilon} t^{-1} \sin(ta) dt \right| \ll \min \{ |a\varepsilon|, 1 \}, \]
\[ \left| \int_{|t| \geq M} t^{-1} \sin(ta) dt \right| \ll \min \left\{ \frac{1}{\alpha M}, 1 \right\}, \]
we obtain
\[ |I_2| \ll (|\bar{P}_n(x)| + |\bar{P}_n(y)| + |\bar{Q}(x)| + |\bar{Q}(y)|) A_x^{-\alpha \rho/6} + \min \{ 1, |\bar{P}_n(x)|^{-1} A_x^{-\alpha \rho/3} \} \]
\[ + \min \{ 1, |\bar{P}_n(y)|^{-1} A_x^{-\alpha \rho/3} \} + \min \{ 1, |\bar{Q}(x)|^{-1} A_x^{-\alpha \rho/3} \} + \min \{ 1, |\bar{Q}(y)|^{-1} A_x^{-\alpha \rho/3} \}. \]

From this and the Gaussianity of \( \bar{P}_n \) and \( \bar{Q} \), we have
\[ \mathbb{E} I_2^2 \ll A_x^{-\alpha \rho/3} + \mathbb{E} \min \{ 1, Z^{-2} A_x^{-2\alpha \rho/3} \} \ll A_x^{-\alpha \rho/3} \]
where \( Z \sim \mathcal{N}(0,1) \).

For \( I_1 \), we need to make use of the cancellation between \( \bar{P}_n \) and \( \bar{Q} \). We rewrite \( I_1 \) as
\[ I_1 = \frac{1}{\pi^2} \int_{A_x^{-\alpha \rho/6}}^{A_x^{\alpha \rho/3}} \int_{A_x^{-\alpha \rho/6}}^{A_x^{\alpha \rho/3}} t^{-1} u^{-1} \sin(tQ(x)) \cos \left( \frac{u}{2} \bar{Q}(y) + \frac{P_n(y)}{2} \right) \sin \left( \frac{u}{2} \bar{Q}(y) - \bar{P}_n(y) \right) dt du \]
\[ + \frac{1}{\pi^2} \int_{A_x^{-\alpha \rho/6}}^{A_x^{\alpha \rho/3}} \int_{A_x^{-\alpha \rho/6}}^{A_x^{\alpha \rho/3}} t^{-1} u^{-1} \sin(u\bar{P}_n(y)) \cos \left( \frac{t}{2} \bar{Q}(x) + \frac{\bar{P}_n(x)}{2} \right) \sin \left( \frac{t}{2} \bar{Q}(x) - \bar{P}_n(x) \right) dt du. \]

Using \( \int_{b}^{c} t^{-1} \sin(ta) dt \ll 1 \) for all \( 0 < b < c, |\frac{\sin(a)}{a}| \leq 1 \) for all \( a \neq 0 \), and (62), we get
\[ \mathbb{E} I_1^2 \ll \mathbb{E} \left[ \int_{A_x^{-\alpha \rho/6}}^{A_x^{\alpha \rho/3}} |Q(y) - \bar{P}_n(y)| + |Q(x) - \bar{P}_n(x)| dt \right]^2 \ll A_x^{2\alpha \rho/3} (\mathbb{E}|Q(y) - \bar{P}_n(y)|^2 + \mathbb{E}|Q(x) - \bar{P}_n(x)|^2) \ll A_x^{-\alpha \rho/3} \]
where we used Lemma 6.2 (recalling that the random variables \( \xi_i \) are iid standard Gaussian and hence have mean 0) to get

\[
\mathbb{E} |\bar{Q}(y) - \bar{P}_n(y)|^2 = \text{Var} (\bar{Q}(y) - \bar{P}_n(y)) \ll A_y^{-\alpha'}/2 \ll A_x^{-\alpha'}/2.
\]

This completes the proof of Lemma 6.3.

\[ \square \]

**Proof of Lemma 6.3** Let \( C_0 \) be the constant in Lemma 5.1.

**Case 1.** \( T \leq 1 \). Since \( T \gg 1 \), it suffices to show that

\[
(76) \quad \mathbb{E}(\epsilon_{a,b,\delta}^{\text{trun}})^4 \ll 1.
\]

For simplicity, we write \( S_{a,b,\delta}^{\text{trun}} \) for \( S_{a,b,\delta}^{\text{sign}} \). Let \( S_{a,b,\delta}^{\text{sign}} = S_{a,b,\delta}^{\text{sign}} \) as in the setting of Lemma 6.4. By the definition of sign changes, we have with probability 1,

\[
S_{a,b,\delta}^{\text{trun}} \ll j_1 - j_0 \ll \delta^{-1} \quad \text{and} \quad S_{a,b,\delta}^{\text{sign}} \ll j_1 - j_0 \ll \delta^{-1}.
\]

Hence, by Lemma 6.4 Hölder’s inequality, and the assumption that \( \delta \geq (\log \frac{1}{a_n})^{-\alpha'}/24 \), we have

\[
|\mathbb{E}(S_{a,b,\delta}^{\text{trun}})^4 - \mathbb{E}(S_{a,b,\delta}^{\text{sign}})^4| \ll \delta^{-3} \mathbb{E}|S_{a,b,\delta}^{\text{trun}} - S_{a,b,\delta}^{\text{sign}}| \ll \delta^{-4} \left( \log \frac{1}{a_n} \right)^{-\alpha'/6} \ll 1.
\]

Thus, it suffices to show that \( \mathbb{E}(S_{a,b,\delta}^{\text{sign}})^4 \ll 1 \). Since \( N_{a,b}(x,y) \leq N_n(x,y) \) for any interval \((x, y)\),

\[
\mathbb{E}(S_{a,b,\delta}^{\text{sign}})^4 \leq \mathbb{E}N_{a,b}^4(x, y).
\]

Partition the interval \((a, b)\) into smaller intervals \((x, y)\) such that \( \log \frac{1-x}{1-y} = \frac{1}{2C_0} \). Since \( \log \frac{1-a}{1-b} = T \), the number of such sub-intervals is \( 2C_0T \). By (45), for each of these intervals \((x, y)\), we have

\[
\mathbb{E}N_{a,b}^4(x, y) \ll \sum_{k=1}^{\infty} k^4 2^{-k/2} \ll 1.
\]

Using this and the assumption that \( T \leq 1 \) of Case 1, we have \( \mathbb{E}N_{a,b}^4(x, y) \ll 1 \) as desired.

**Case 2.** \( T > 1 \). We decompose the sum in \( S_{a,b,\delta}^{\text{trun}} - \mathbb{E}S_{a,b,\delta}^{\text{trun}} \) into blocks of size \( \mu := \delta^{-1} \) of the form

\[
X_k = \sum_{j=(j_0+(k-1)\mu)}^{j_0+k\mu-1} (N_{a,b,\delta}^{\text{trun}}(x_j, x_{j+1}) - \mathbb{E}N_{a,b,\delta}^{\text{trun}}(x_j, x_{j+1}))
\]

for each \( k = 1, \ldots, j_2 \), where \( j_2 = (j_1 - j_0)\mu^{-1} \) is the number of blocks. Notice that \( j_2 \ll T \).

We have

\[
\mathbb{E} (S_{a,b,\delta}^{\text{trun}} - \mathbb{E}S_{a,b,\delta}^{\text{trun}})^4 = \mathbb{E} \left( \sum_{k=1}^{j_2} X_k \right)^4 = \sum_{k=1}^{j_2} \mathbb{E}X_k^4 + 4 \sum_{k \neq l} \mathbb{E}X_k^3 X_l + 6 \sum_{k < l} \mathbb{E}X_k^2 X_l^2 + 12 \sum_{l < p, k \neq l, p} \mathbb{E}X_k^2 X_l X_p + 24 \sum_{k < l < p < q} \mathbb{E}X_k X_l X_p X_q
\]

\[
= I_1 + 4I_2 + 6I_3 + 12I_4 + 24I_5.
\]

Note that each \( X_k \) is of the form \( S_{a',b',\delta}^{\text{trun}} - \mathbb{E}S_{a',b',\delta}^{\text{trun}} \) for some \( a', b' \) that satisfy \( \log \frac{1-a'}{1-b'} \leq 1 \). Thus, (76) implies that \( \mathbb{E}X_k^4 \ll 1 \) for all \( k \). By Hölder’s inequality each term in the summation of \( I_1, \ldots, I_5 \) is of order \( O(1) \) and so,

\[
I_1 \ll j_2 \ll T, \quad I_2 + I_3 \ll j_2^2 \ll T^2.
\]
To bound $I_4$ and $I_5$, we use the independence in Lemma 6.2 to conclude that if $k_2 - k_1 \geq 3\alpha \log \log n$ then $X_{k_2}$ and $(X_1, \ldots, X_{k_1})$ are independent. Together with the fact that $\mathbb{E}X_k = 0$ for all $k$, we observe that most terms in the sums $I_4, I_5$ are zero. Ignoring these terms, we have

$$I_4 = \sum_{l < p \leq l + C \log \log n} \mathbb{E}X_k^2X_lX_p \ll j_2^2 \log \log n \ll T^2 \log \log n,$$

and

$$I_5 = \sum_{l - C \log \log n \leq k < l < p < q \leq p + C \log \log n} \mathbb{E}X_kX_lX_pX_q \ll T^2(\log \log n)^2.$$

Putting the above bounds together, we obtain Lemma 6.5.

\[ \square \]

7. PROOF OF LEMMA 2.5

Since in this section we only deal with Gaussian random variables, we again use $\xi_i$ to denote iid standard Gaussian variables (instead of $\tilde{\xi}_i$). This would help avoid complicated notation (such as double superscripts) later on. By symmetry of the Gaussian distribution, we can assume that $c_i \geq 0$ for all $i$.

Let

$$a_n = \exp \left( -2 \log^{1/5} n \right) \quad \text{and} \quad b_n = \frac{1}{a_n n}$$

and

$$\mathcal{J} := \pm (1 - a_n, 1 - b_n) \cup \pm (1 - a_n, 1 - b_n)^{-1}.$$  

Note that this $a_n$ satisfies Condition (2).

By Proposition 2.3,

$$\mathbb{E}N_n^2(\mathbb{R} \setminus \mathcal{J}) \ll \log^4 \frac{1}{a_n} = o(\log n).$$

Thus, to prove Lemma 2.5, it suffices to show that

$$\text{Var } N_{P_n}(\mathcal{J}) = \Omega(\log n).$$

We have

$$N_{P_n}(\mathcal{J}) = N_{P_n}(\mathcal{J} \cap [-1, 1]) + N_{P_n}(\mathcal{J} \setminus [-1, 1])$$

= $$N_{P_n}(\mathcal{J} \cap [-1, 1]) + N_{R_n}(\mathcal{J} \setminus [-1, 1]),$$

where $R_n(x) = \frac{c_n x}{1 - c_n x} = \sum_{i=0}^{n} \frac{c_n^{i-1} c_n x^i}{c_n^{i-1} x^i}$.

Since $\text{Var } (X + Y) = \text{Var } X + \text{Var } Y + \text{Cov } (X, Y) \geq \text{Var } X + \text{Cov } (X, Y)$ for any two real random variables $X$ and $Y$, it suffices to show that

$$\text{Var } N_{R_n}(\mathcal{J} \cap [-1, 1]) = \Omega(\log n)$$

and

$$\text{Cov } (N_{P_n}(\mathcal{J} \cap [-1, 1]), N_{R_n}(\mathcal{J} \cap [-1, 1])) = o(\log n).$$

In order to verify (78), we use the universality method in a novel way. Instead of swapping the random variables $\xi_i$, we swap the deterministic coefficients $c_i$. This allows us to couple $R_n$ with the Kac polynomial and the desired bound follows by known results concerning the variance of the Kac polynomial. This swapping is possible thanks to the fact that the “important” coefficients are $\frac{c_n^{i-1}}{c_n}$ which are close to 1 by (1).
Let 
\[ \tilde{R}_n(x) := \sum_{i=0}^{n} \xi_{n-i} x^i \]
be the corresponding Kac polynomial. We prove the following analogs of Theorem 2.1 and Corollary 2.2 for \( R_n \) and \( \tilde{R}_n \).

**Proposition 7.1.** Assume that the \( \xi_i \) are iid standard Gaussian. Let \( \beta > 0 \) be any constant. There exists a constant \( C > 0 \) such that for every function \( F : \mathbb{R} \to \mathbb{R} \) whose derivative up to order 3 are bounded by 1 and for every \( n \), we have
\[
\left| E F \left( N_{R_n} (J \cap [-1, 1]) \right) - E F \left( N_{\tilde{R}_n} (J \cap [-1, 1]) \right) \right| \leq C (\log n)^{-\beta}.
\]

**Proposition 7.2.** Assume that the \( \xi_i \) are iid standard Gaussian. Let \( \beta > 0 \) be any constant. There exists a constant \( C > 0 \) such that for every \( n \), we have
\[
\left| E \left( N_{R_n}^k (J \cap [-1, 1]) \right) - E \left( N_{\tilde{R}_n}^k (J \cap [-1, 1]) \right) \right| \leq C (\log n)^{-\beta}
\]
for \( k = 1, 2 \). In particular, 
\[
\left| \text{Var} \left( N_{R_n} (J \cap [-1, 1]) \right) - \text{Var} \left( N_{\tilde{R}_n} (J \cap [-1, 1]) \right) \right| \leq C (\log n)^{-\beta}.
\]

It is easy to show that Proposition 7.1 implies Proposition 7.2 using the same arguments as in the proof of Corollary 2.2. Assuming Proposition 7.1, we next prove (78) and (79).

**Proof of (78).** As shown in [24], for the Kac polynomial \( \tilde{R}_n \) (recall that the random variables \( \xi_i \) are iid standard Gaussian), \( \text{Var} \left( N_{\tilde{R}_n} ((-1, 1)) \right) \gg \log n \).

By Proposition 2.3 for the Kac polynomial \( \tilde{R}_n \) and the choice of \( a_n, b_n \) in (77),
\[ EN_{\tilde{R}_n}^2 ([1, 1] \cap J) \ll EN_{\tilde{R}_n}^2 (\mathbb{R} \setminus J) = o (\log n). \]

So, by the triangle inequality,
\[ \sqrt{\text{Var} \left( N_{\tilde{R}_n} (J \cap [-1, 1]) \right)} \geq \sqrt{\text{Var} \left( N_{\tilde{R}_n} ((-1, 1)) \right) - o(\sqrt{\log n})} \gg \sqrt{\log n}. \]

This together with Proposition 7.2 imply (78).

**Proof of (79).** By a classical formula [20, Theorem 1], we have that for every \( a < b \) and for every nonzero polynomial \( f \),
\[
N_f(a, b) = \frac{1}{2\pi} \int_a^b |f'(x)| \cos(sf(x))duds = \frac{1}{2\pi^2} \int_a^b \int_a^b \frac{1}{u^2} (1 - \cos(uf'(x))) \cos(sf(x))duds.
\]

We will apply this formula for both \( P_n \) and \( R_n \). To avoid the improper integrals, we need to cut off the domain of integration. Let \( D := \exp(a_n^{-1}/100) \), \( \gamma := D^{-3} \) and approximate \( N_f(a, b) \) by
\[
N_f^{(1)}(a, b) := \frac{1}{2\pi} \int_{-D}^D \int_a^b |f'(x)| \cos(sf(x))duds,
\]
and
\[
N_f^{(2)}(a, b) := \frac{1}{2\pi^2} \int_{-D}^D \int_a^b \int_{-D}^D \frac{1}{u^2} (1 - \cos(uf'(x))) \cos(sf(x))duds.
\]
We first show that $N_f^{(1)}$ is a good approximation of $N_f$. We claim that for any $(a, b) \subset \mathbb{J} \cap [-1, 1]$,

$$
\mathbb{E} \left[ |N_P_n(a, b) - N_P^{(1)}(a, b)|^2 \right] \ll \exp(-\Omega(a_n^{-1})) \, .
$$

To show this, let $x_1 < \cdots < x_k$ be all the roots of $P'_n(x)$ in the interval $(a, b)$ and let $x_0 = a, x_{k+1} = b$. We have $k \leq n$. Since $P_n$ keeps the same sign on each interval $(x_i, x_{i+1})$, we have

$$
N_P_n(a, b) - N_P^{(1)}(a, b) \leq \frac{1}{2\pi} \sum_{i=0}^{k} \left| \int_{|s| \geq D} \frac{\sin(sP_n(x_{i+1})) - \sin(sP_n(x_i))}{s} ds \right|
$$

$$
\ll \sum_{i=0}^{k+1} \min \left\{ 1, \frac{1}{|D|P_n(x_i)|} \right\}
$$

where we used (75). Thus,

$$
\mathbb{E} \left[ |N_P_n(a, b) - N_P^{(1)}(a, b)|^2 \right]^{1/2} \ll \sum_{i=0}^{k+1} \left( \mathbb{E} \min \left\{ 1, \frac{1}{|D|P_n(x_i)|} \right\} \right)^{1/2} \, .
$$

We divide the interval $(a, b)$ into $D^{1/2}$ equal intervals by the points $a = a_0 < a_1 < \cdots < a_{D^{1/2}} = b$. Let $p = 1/4$ (or any small constant). For each $1 \leq i \leq k$, assume that $x_i \in (a_j, a_{j+1}]$ for some $j$. If $|P_n(x_i)| \leq D^{p-1}$ and $|P_n(a_{j+1})| \geq 2D^{p-1}$ then

$$
|P_n(a_{j+1}) - P_n(x_i)| = \left| \int_{x_i}^{a_{j+1}} \int_{x_i}^{t} P''_n(u) du dt \right| \geq D^{p-1}
$$

and so

$$
\int_{a_j}^{a_{j+1}} \int_{a_j}^{a_{j+1}} |P''_n(u)| du dt \geq D^{p-1} \, .
$$

We claim that this happens with small probability

$$
\mathbb{P} \left( \int_{a_j}^{a_{j+1}} \int_{a_j}^{a_{j+1}} |P''_n(u)| du dt \geq D^{p-1} \right) \ll D^{-1} \, .
$$

We defer the proof of (82) to Appendix 8.7 as it is similar to the proof of Lemma 5.1. Using this and the union bound over all $D^{1/2}$ possible values of $j$, we get

$$
P \left( |P_n(x_i)| \leq D^{p-1} \right) \leq \mathbb{P} \left( \exists j : |P_n(a_j)| \leq 2D^{p-1} \right) + D^{1/2}O(D^{-1}) \ll D^{1/2}D^{p-1} + D^{-1/2} \ll D^{-1/4}
$$

where we used the fact that $P_n(a_j)$ is a Gaussian random variable with variance $\Omega(1)$. Plugging this into (81) and using $k \leq n, p = 1/4$, we obtain

$$
\left( \mathbb{E} \left[ |N_P_n(a, b) - N_P^{(1)}(a, b)|^2 \right] \right)^{1/2} \ll n \left( D^{-1/4} + D^{-1/8} \right) \ll \exp(-\Omega(a_n^{-1})) \, .
$$

This proves (80), which means that $N_P^{(1)}(a, b)$ is a good approximation of $N_P(a, b)$.

Next, we show that for all $(a, b) \subset \mathbb{J} \cap [-1, 1]$, $N_f^{(2)}(a, b)$ is also a good approximation of $N_P^{(1)}(a, b)$, namely,

$$
\mathbb{E} \left[ |N_P^{(1)}(a, b) - N_P^{(2)}(a, b)|^2 \right] \ll \exp(-\Omega(a_n^{-1})) \, .
$$
To start, using the fact that $0 \leq 1 - \cos x \leq x^2$ for every real number $x$, we have
\[
\left| N_{P_n}^{(1)}(a, b) - N_{P_n}^{(2)}(a, b) \right| \ll \int_{-D}^{D} \int_{a}^{b} \int_{0}^{\gamma} |P'_n(x)|^2 du dx ds + \int_{-D}^{D} \int_{a}^{b} \int_{D^2}^{\infty} \frac{1}{u^2} du dx ds \\
\ll D^{-2} \int_{a}^{b} |P'_n(x)|^2 dx + D^{-1}.
\]
Taking the second moment of both sides, we get
\[
\mathbb{E} \left| N_{P_n}^{(1)}(a, b) - N_{P_n}^{(2)}(a, b) \right|^2 \ll D^{-1} + D^{-2} \int_{a}^{b} \mathbb{E} |P'_n(x)|^4 dx \ll D^{-1} + D^{-2} n O(1) \ll \exp \left( -\Omega(a_n^{-1}) \right)
\]
where we again used the fact that $a_n$ satisfies (2). This proves (83).

Combining this with (80), we conclude that for any $(a, b) \subset \mathcal{I} \cap [-1, 1],
\[
\mathbb{E} \left| N_{P_n}(a, b) - N_{P_n}^{(2)}(a, b) \right|^2 \ll \exp \left( -\Omega(a_n^{-1}) \right).
\]

We can obtain a similar estimate for $R_n$. Therefore, in order to prove (79), it suffices to show
\[
\text{Cov} \left( N_{P_n}^{(2)}(\mathcal{I} \cap [-1, 1]), N_{R_n}^{(2)}(\mathcal{I} \cap [-1, 1]) \right) = o(\log n).
\]
To prove this bound, we need to make a critical use of a property of Gaussian variable. For a standard Gaussian random variable $Z$ and any real number $a$, $\mathbb{E} \cos(aZ) = E e^{iaZ} = e^{-a^2/2}$. Since $P_n(x), R_n(x)$ are Gaussian for any value of $x$, we have for $(a, b), (c, d) \subset \mathcal{I} \cap [-1, 1],
\]
\[
\text{Cov} \left( N_{P_n}^{(2)}(a, b), N_{R_n}^{(2)}(c, d) \right) = \frac{1}{4\pi^4} \int_{a}^{b} \int_{c}^{d} \int_{\gamma}^{D^2} \int_{-D}^{D} \int_{-D}^{D} \frac{1}{u^2v^2} (F_1 + F_2 + F_3 + F_4) dt ds dv du dx
\]
where
\[
F_1(x, y, u, v, s, t) := \frac{1}{8} \sum \exp \left( -\frac{1}{2} \text{Var} \left( sP_n(x) \pm uP'_n(x) \pm tR_n(y) \pm vR'_n(y) \right) \right)
\]
\[
- \frac{1}{4} \sum \exp \left( -\frac{1}{2} \text{Var} \left( sP_n(x) \pm uP'_n(x) \right) - \frac{1}{2} \text{Var} \left( tR_n(y) \pm vR'_n(y) \right) \right)
\]
in which the sums are taken over all possible assignments of $+$ and $-$ signs in place of the $\pm$ and
\[
F_2(x, y, u, v, s, t) := -F_1(x, y, 0, v, s, t),
F_3(x, y, u, v, s, t) := -F_1(x, y, u, 0, s, t),
F_4(x, y, u, v, s, t) := F_1(x, y, 0, 0, s, t).
\]

These formulas follow directly from the definition of $N^{(2)}$; we provide the tedious derivation in Appendix 8.8 for the reader’s convenience.

We now show that for $(a, b), (c, d) \subset \mathcal{I} \cap [-1, 1]$ and for all $i = 1, 2, 3, 4,$
\[
\int_{a}^{b} \int_{c}^{d} \int_{\gamma}^{D^2} \int_{-D}^{D} \int_{-D}^{D} \frac{1}{u^2v^2} F_i dt ds dv du dx = o(1).
\]
We will show it for $i = 4$. The cases $i = 1, 2, 3$ are completely similar. We have
\[
F_4(x, y, u, v, s, t) = \exp \left( -\frac{s^2}{2} \sum_{i=0}^{n} c_i^2 x^{2i} \right) \exp \left( -\frac{t^2}{2} \sum_{i=0}^{n} \frac{c_i^2 y^{2i}}{c_n^2} \right) \left( \frac{e^{st\Delta} + e^{-st\Delta}}{2} - 1 \right)
\]
where
\[
\Delta = \sum_{i=0}^{n} \frac{c_i^2}{c_n^2} x^{2i} y^{n-i}.
\]
Since $|x|, |y| \leq 1 - b_n$ and $nb_n \geq a_n^{-1}/2 \gg C \log n$ for any constant $C$, we have
\[
\Delta \ll n^{O(1)} \sum_{i=0}^{n} (1 - b_n)^n \ll n^{O(1)} \exp (-nb_n) \ll \exp (-a_n^{-1}/4).
\]
Thus, bounding the first two exponents in (87) by 1, using $D = \exp \left( a_n^{-1}/100 \right)$ and $|s|, |t| \leq D$, we get that on the domain of integration in (86),
\[
F_4(x, y, u, v, s, t) \ll \exp \left( O(1) D^2 \exp \left( -a_n^{-1}/4 \right) \right) - 1 \ll D^2 \exp \left( -a_n^{-1}/4 \right) \ll \exp \left( -a_n^{-1}/5 \right).
\]
Finally, using $\gamma = D^{-3}$, we have
\[
\int_a^b \int_c^d \int_{D^2} \int_{D^2} \int_{-D}^D \int_{-D}^D \frac{1}{u^2v^2} F_4 dt dv dx dy \ll D^8 \exp \left( -a_n^{-1}/5 \right) = o(1)
\]
proving (86) and completing the proof of (79).

Proof of Proposition 7.1. We use the same arguments as in the proof of Theorem 2.1 with the following modifications. First, $P_n$ is of course, replaced by $R_n$ and $\tilde{P}_n$ is replaced by $\tilde{R}_n$, and all of the $\delta^\alpha$ in the former for a small constant $\alpha$ with be replaced by $(\log n)^{-\beta}$ for a large constant $\beta$. For example, Lemma 3.4 is replaced by the following variant that can be proved using the same argument.

Lemma 7.3. Assume that the $\xi_i$ are iid standard Gaussian. Let $\delta \in [b_n, a_n]$. For any constant $\gamma > 0$ and $x \in \mathbb{R}$ with $|x| \in [1 - \delta - \delta(\log n)^{-\gamma}, 1 - \delta/2 + \delta(\log n)^{-\gamma}]$, we have
\[
\mathbb{P} \left( N_{R_n}(x, \delta(\log n)^{-\gamma}) \geq 2 \right) \ll (\log n)^{-3\gamma/2}.
\]

The only remaining difference compared to the proof of Theorem 2.1 is in the proof of the analog of Lemma 3.7, namely for $\delta_0 = a_n, \delta_1 = a_n/2, \ldots, \delta_{M-1} = a_n/2^{M-1}$ and $\delta_M := \max \{1/n, b_n\}$ ($M$ is the largest integer such that $\delta_{M-1} > \max \{1/n, b_n\}$), and for $m_i = (\log n)^{\beta_i}$,

Lemma 7.4. Assume that the $\xi_i$ are iid standard Gaussian. Let $\beta'$ be any positive constant. Let $L : \mathbb{C}^{m_1 + \cdots + m_M} \to \mathbb{R}$ be a smooth function with all derivatives up to order 3 being bounded by $(\log n)^{-\beta'}$. Then for every $w_{ik}$ in $B(1 - 3\delta_i/2, 2\delta_i/3)$, we have
\[
\left| E \left( \frac{R_n(w_{ik})}{\sqrt{V(w_{ik})}} \right) \right|_{i=1,\ldots,M} - \left| E \left( \frac{\tilde{R}_n(w_{ik})}{\sqrt{V(w_{ik})}} \right) \right|_{i=1,\ldots,M} \ll (\log n)^{-\beta'},
\]
where $V(w) := \text{Var } R_n(w)$.

Assuming this lemma, the rest of the proof of Theorem 2.1 can be adapted in a straightforward manner to complete the proof of Proposition 7.1. □
Proof of Lemma 7.4. While for Lemma 3.7, going from $P_n$ to $\tilde{P}_n$, we need to swap the general random variables $ξ_i$ to the Gaussian ones $ξ_i$, here, going from $R_n$ to $\tilde{R}_n$, we need to swap the coefficients $\frac{c_{n-i}}{c_n}$ to 1 and keep the Gaussian random variables $ξ_i$ intact. Keeping that in mind, we set for each $0 \leq i_0 \leq n + 1$,

$$R_{i_0}(z) := \sum_{i=0}^{i_0-1} \xi_{n-i}z^i + \sum_{i=i_0}^n \frac{c_{n-i}}{c_n} \xi_{n-i}z^i.$$ 

We have $R_0 = R_n$, $R_{n+1} = \tilde{R}_n$ and $R_{i_0+1}$ is obtained from $R_{i_0}$ by replacing the coefficient $\frac{c_{n-i_0}}{c_n}$ by 1.

The difference $d_{i_0}$ in (28) for $0 \leq i_0 \leq n$ now becomes

$$d_{i_0} := \left| \mathbb{E}_{ξ_{n-i_0}} \hat{L} \left( \frac{c_{n-i_0}ξ_{n-i_0}w_{ik}^{i_0}}{c_n \sqrt{V(w_{ik})}} \right) \right|_{ik} - \mathbb{E}_{ξ_{n-i_0}} \hat{L} \left( \frac{ξ_{n-i_0}w_{ik}^{i_0}}{\sqrt{V(w_{ik})}} \right)_{ik}$$

where $\hat{L}$ is obtained from $L$ by translation and thus has all derivatives up to order 3 bounded by $(\log n)^{3\beta'\gamma}$. The task is to show that

$$\sum_{i_0=0}^{n+1} \mathbb{E}_{ξ_0,...,ξ_n} d_{i_0} \ll (\log n)^{-\beta'}. $$

By the Taylor expansion of order 2, we get

$$\hat{L} \left( \frac{c_{n-i_0}ξ_{n-i_0}w_{ik}^{i_0}}{c_n \sqrt{V(w_{ik})}} \right)_{ik} = \hat{L}(0) + \frac{1}{2} \hat{L}_2 + \text{err}_2,$$

where

$$\hat{L}_1 := \left. \frac{d \hat{L} \left( \frac{c_{n-i_0}ξ_{n-i_0}w_{ik}^{i_0}}{c_n \sqrt{V(w_{ik})}} \right)_{ik}}{dt} \right|_{t=0}$$

$$= \sum_{ik} \frac{\partial \hat{L}(0)}{\partial \text{Re}(z_{ik})} \text{Re} \left( \frac{c_{n-i_0}ξ_{n-i_0}w_{ik}^{i_0}}{c_n \sqrt{V(w_{ik})}} \right) + \sum_{ik} \frac{\partial \hat{L}(0)}{\partial \text{Im}(z_{ik})} \text{Im} \left( \frac{c_{n-i_0}ξ_{n-i_0}w_{ik}^{i_0}}{c_n \sqrt{V(w_{ik})}} \right),$$

$$\hat{L}_2 := \left. \frac{d^2 \hat{L} \left( \frac{c_{n-i_0}ξ_{n-i_0}w_{ik}^{i_0}}{c_n \sqrt{V(w_{ik})}} \right)_{ik}}{dt^2} \right|_{t=0},$$

and

$$|\text{err}_2| \ll (\log n)^{3\beta'\gamma} \frac{c_n}{c_0} \frac{|ξ_{n-i_0}|^3}{|w_{ik}|^{i_0}} \left( \sum_{ik} |w_{ik}|^{i_0} \right)^{1/2} \left( \sum_{i_0} |w_{ik}|^{i_0} \right)^{1/2} \delta_i^{-1/2} \delta_i^{-1/2} \delta_i^{-1/2} \delta_i^{-1/2} \delta_i^{-1/2}.$$
Similarly, we get the expansion for $\hat{L}_i \left( \frac{\xi_{n-i_0}w_{i_k}}{\sqrt{V(w_{i_k})}} \right)$. Subtracting the two expansions and taking expectation both sides (noting again that all of the $\xi_i$ are iid standard Gaussian and in particular, have mean 0 and variance 1), we obtain

$$
(\log n)^{-3\beta'} d_{i_0} \ll \left| \frac{c_{n,i_0}^2}{c_{n}^2} - 1 \right| \left( \sum_{ik} |w_{i_k}|^{10}\delta_{i_0}^{1/2} \right)^2 + \left( \frac{c_{n,i_0}^3}{c_{n}^3} + 1 \right) \left( \sum_{ik} |w_{i_k}|^{10}\delta_{i_0}^{1/2} \right)^3
$$

(92)

$$
\ll (\log n)^{O(\beta')} \left| \frac{c_{n,i_0}^2}{c_{n}^2} - 1 \right| \sum_{i=1}^M \delta_i (1 - \delta_i/2)^{2i_0} + (\log n)^{O(\beta')} \left( \frac{c_{n,i_0}^3}{c_{n}^3} + 1 \right) \sum_{i=1}^M \delta_i (1 - \delta_i/2)^{2i_0}
$$

where in the last inequality, we used $|w_{i_k}| \leq 1 - \delta_i/2$, Cauchy-Schwarz inequality and the fact that $M \ll \log n$. Note that for each $i$,

$$
\sum_{i_0=0}^{n} c_{n,i_0}^3 (1 - \delta_i/2)^{2i_0} \leq \sum_{i_0=0}^{n/2} c_{n,i_0}^3 (1 - \delta_i/2)^{2i_0} + \sum_{i_0=n/2}^{n} c_{n,i_0}^3 (1 - \delta_i/2)^{2i_0}
$$

$$
\ll \sum_{i_0=0}^{n/2} (1 - \delta_i/2)^{2i_0} + n^{O(1)} (1 - \delta_i/2)^n \ll \delta_i^{-1}
$$

Thus, plugging this into (92) and using $\sum_{i=1}^M \delta_i^{1/2} \ll \delta_i^{1/2} \ll (\log n)^{-C}$ for any constant $C$,

$$
(\log n)^{-3\beta'} \sum_{i_0=0}^{n} d_{i_0} \ll (\log n)^{O(\beta')} \sum_{i=1}^M \sum_{i_0=0}^{n} \left| \frac{c_{n,i_0}^2}{c_{n}^2} - 1 \right| \delta_i (1 - \delta_i/2)^{2i_0} + (\log n)^{O(\beta')} \sum_{i=1}^M \delta_i^{1/2}
$$

Let

$$
I_0 := a_n^{-1/2} = \exp \left( \log 1/5 \right) n \quad \text{and} \quad I_1 := \frac{(\log n)^2}{b_n} \leq n \exp \left( -\log 1/5 \right) n.
$$

Splitting the double sum

$$
\sum_{i=1}^M \sum_{i_0=0}^{n} \left| \frac{c_{n,i_0}^2}{c_{n}^2} - 1 \right| \delta_i (1 - \delta_i/2)^{2i_0}
$$

into $\sum_{i=1}^M \sum_{i_0=I_0}^{I_1} \sum_{i=1}^M \sum_{i_0=0}^{I_0-1} \sum_{i=0}^{I_1} \sum_{i_0=I_1+1}^{n}$ and denoting the corresponding sums by $S_1, S_2, S_3$, we obtain

$$
(\log n)^{-3\beta'} \sum_{i_0=0}^{n} d_{i_0} \ll (\log n)^{O(\beta')} (S_1 + S_2 + S_3) + (\log n)^{-4\beta'}.
$$

By assumption [1], we have for every $i_0 \in [I_0, I_1]$,

$$
\frac{2}{c_{n,i_0}^2} - 1 \ll \exp \left( -\log (\log n)^{1+\epsilon} \right).
$$

Hence,

$$
S_1 \ll \exp \left( -\log (\log n)^{1+\epsilon} \right) \sum_{i=1}^M \sum_{i_0=0}^{n} \delta_i (1 - \delta_i/2)^{2i_0} \ll M \exp \left( -\log (\log n)^{1+\epsilon} \right).
$$
For $S_2$, we observe that $\frac{c^2_{n-i_0}}{a_n^2} \ll 1$ for all $i_0 \leq I_0 \leq n/2$ by Condition [A3] and so

$$S_2 \ll \sum_{i=1}^{M} \sum_{i_0=0}^{I_0-1} \delta_i \ll I_0 a_n = a_n^{1/2}.$$ 

For $S_3$, we observe that $\frac{c^2_{n-i_0}}{a_n^2} \ll n^{O(1)}$ for all $i_0$ by Condition [A3] and that for all $i_0 \geq I_1$,

$$(1 - \delta_i/2)^{i_0} \ll (1 - b_n/2)^{I_1} \ll \exp (-b_n I_1/2) = \exp(-\log^2 n).$$

And so,

$$S_3 \ll n^{O(1)} \sum_{i=1}^{M} \sum_{i_0=I_1+1}^{n} \exp(-\log^2 n) \ll n^{O(1)} \exp(-\log^2 n).$$

Combining these bounds, we obtain

$$(\log n)^{-3\beta'} \sum_{i_0=0}^{n} d_{i_0} \ll (\log n)^{-3\beta'}$$

proving (90) and completing the proof of Lemma 7.4. \hfill \Box

**Acknowledgements.** O. Nguyen would like to thank François Baccelli and Terence Tao for helpful suggestions.

**REFERENCES**

[1] Jean-Marc Azaïs, Federico Dalmao, and José R León. CLT for the zeros of classical random trigonometric polynomials. In *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, volume 52, pages 804–820. Institut Henri Poincaré, 2016.

[2] Jean-Marc Azaïs and José León. CLT for crossings of random trigonometric polynomials. *Electronic Journal of Probability*, 18, 2013.

[3] Albert T Bharucha-Reid and M Sambandham. *Random Polynomials: Probability and Mathematical Statistics: a Series of Monographs and Textbooks*. Academic Press, 1986.

[4] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[5] Federico Dalmao. Asymptotic variance and CLT for the number of zeros of Kostlan Shub Smale random polynomials. *Comptes Rendus Mathematique*, 353(12):1141–1145, 2015.

[6] Yen Do, Oanh Nguyen, and Van Vu. Roots of random polynomials with coefficients with polynomial growth. *Annals of probability*, 46(5):2407–2494, 2018.

[7] Yen Do and Van Vu. Central limit theorems for the real zeros of Weyl polynomials. arXiv:1707.09276, to appear in *American Journal of Mathematics*, 2017.

[8] Alan Edelman and Eric Kostlan. How many zeros of a random polynomial are real? *Bulletin of the American Mathematical Society*, 32(1):1–37, 1995.

[9] Paul Erdős and AC Offord. On the number of real roots of a random algebraic equation. *Proceedings of the London Mathematical Society*, 3(1):139–160, 1956.

[10] Kambiz Farahmand. *Topics in random polynomials*, volume 393. CRC Press, 1998.

[11] Hendrik Flasche and Zakhar Kabluchko. Real zeros of random analytic functions associated with geometries of constant curvature. arXiv preprint arXiv:1802.02390, 2018.

[12] Friedrich Götze, Dzianis Kaliada, and Dmitry Zaporozhets. Correlation functions of real zeros of random polynomials. arXiv preprint arXiv:1510.00025, 2015.

[13] Andrew Granville and Igor Wigman. The distribution of the zeros of random trigonometric polynomials. *American journal of mathematics*, 133(2):295–357, 2011.

[14] John Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51. American Mathematical Society Providence, RI, 2009.
8. Appendix

8.1. Proof of equation (15). In this section, we show (15), namely, for $m_i := \delta_i^{-11\alpha}$,

\[
\sum_{j=1}^{n} \varphi_i(\zeta_j) - \frac{2\delta_i^2}{9m_i} \sum_{k=1}^{m_i} \log |P_n(w_{ik})| \Delta \varphi_i(w_{ik}) = O(\delta_i^\alpha)
\]

with probability at least $1 - O(\delta_i^\alpha)$, where $w_{ik}$ are chosen independently, uniformly at random from the ball $B(1 - 3\delta_i/2, 2\delta_i/3)$ and are independent of all random variables $\xi_j$. For notational convenience, we skip the subscript $i$ and write $\delta := \delta_i$, $\varphi := \varphi_i$, and $m := m_i$. Let $x_0 = 1 - 3\delta_i/2$, the center of the ball.
Since \( \varphi \) is compactly supported in \( B(x_0, 2\delta/3) \), by the Green’s second identity, we have
\[
\sum_{j=1}^{n} \varphi(\zeta_j) = \frac{1}{2\pi} \int_{\mathbb{C}} \log |P_n(z)| \cdot |\varphi(z)| \, dz = \frac{1}{2\pi} \int_{B(x_0, 2\delta/3)} \log |P_n(z)| \cdot |\varphi(z)| \, dz.
\]

Let \( T \) be the event on which the following hold for \( c_1 = \alpha/2 \).
\[
(T1) \, \log |P_n(z)| \leq \frac{1}{4} \delta^{-c_1} \text{ for all } z \in B(x_0, \frac{\delta}{3}).
\]
\[
(T2) \, \log |P_n(x_1)| \geq -\frac{1}{2} \delta^{-c_1} \text{ for some } x_1 \in B(x_0, \frac{\delta}{100}).
\]

By Jensen’s inequality [22], these conditions imply
\[
N_n \left( B \left( x_0, \frac{3\delta}{4} \right) \right) \ll \delta^{-c_1}.
\]

We will show later that
\[
P(T) = 1 - O(\delta^\alpha).
\]

Assuming (96), it suffices to show that (93), conditioned on \( T \), holds with probability \( 1 - O(\delta^\alpha) \).

**Lemma 8.1.** On the event \( T \), we have
\[
\int_{B(x_0, 2\delta/3)} (\log |P_n(z)|)^2 \, dz \ll \delta^{-8c_1+2}.
\]

Assuming Lemma 8.1 and the fact that \( \|\varphi\|_\infty \ll \delta^{-2(1+\alpha)} \) by the definition of \( \varphi \), we conclude that on the event \( T \),
\[
\int_{B(x_0, 2\delta/3)} \log^2 |P_n(z)| \cdot |\varphi(z)|^2 \, dz \ll \delta^{-4-8\alpha}
\]
where \( \int_{B(x_0, 2\delta/3)} f(z) \, dz := \frac{1}{|B(x_0, 2\delta/3)|} \int_{B(x_0, 2\delta/3)} f(z) \, dz \) is the average of \( f \) on the domain of integration.

Having bounded the 2-norm, we now use the following sampling lemma which is a direct application of Chebyshev’s inequality.

**Lemma 8.2** (Monte Carlo sampling Lemma). ([37, Lemma 38]) Let \( (X, \mu) \) be a probability space, and \( F : X \to \mathbb{C} \) be a square integrable function. Let \( m \geq 1 \), let \( x_1, \ldots, x_m \) be drawn independently at random from \( X \) with distribution \( \mu \), and let \( S \) be the empirical average
\[
S := \frac{1}{m} (F(x_1) + \cdots + F(x_m)).
\]

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
\[
\mathbb{P} \left( |S - \int_X F \, d\mu| \geq \lambda \right) \leq \frac{1}{m\lambda^2} \int_X \left( F - \int_X F \, d\mu \right)^2 \, d\mu.
\]

Conditioning on \( T \) and applying this sampling lemma with \( \lambda = \delta^{\alpha-2} \) together with (98), we obtain
\[
\int_{B(x_0, 2\delta/3)} \log |P_n(z)| \cdot |\varphi(z)| \, dz - \frac{1}{m} \sum_{k=1}^{m} \log |P_n(w_{ik})| \cdot |\varphi(w_{ik})| \ll \delta^{\alpha-2}.
\]
Proof of Lemma 8.1. Since \(-\frac{\delta^{-4} - s_0}{m^2 m^{-4}} = 1 - \delta^\alpha\) where we recall that \(m = \delta^{-11\alpha}\).

Combining this with (94) gives (93), conditioned on \(T\) as claimed. \(\square\)

Proof of (96). Since

\[
\sup_{z \in B(x_0, \delta/5)} |P_n(z)| \leq \sum_{i=0}^{n} |c_i| |\xi_i| (1 - 7\delta/10)^i
\]

has mean at most \(\delta^{-O(1)}\), applying Markov’s inequality to the random variable \(\sum_{i=0}^{n} |c_i| |\xi_i| (1 - 7\delta/10)^i\), we conclude that the event (1) happens with probability at least 1 \(-\frac{\delta^{-4} - s_0}{m^2 m^{-4}}\).

For (2), writing \(x = re^{i\theta}\) and observing that the set \(\{w = re^{i\theta'} : \theta' \in [\theta - \delta/100, \theta + \delta/100]\}\) is a subset of \(B(x, \delta/100)\), we have

\[
\mathbb{P}(\text{[2] fails}) \leq \mathbb{P}\left(\sup_{\theta' \in [\theta - \delta/100, \theta + \delta/100]} \left| \sum_{j=0}^{n} c_j \xi_j r^j e^{ij\theta'} \right| \leq \exp\left(-\delta^{-c_1}/2\right) \right).
\]

By taking the supremum outside, the right-hand side is at most

\[
\sup_{\theta' \in [\theta - \delta/100, \theta + \delta/100]} \mathbb{P}\left( \left| \sum_{j=0}^{n} c_j \xi_j r^j e^{ij\theta'} \right| \leq \exp\left(-\delta^{-c_1}/2\right) \right)
\]

and hence by projecting onto the real line and conditioning on the random variables \((\xi_j)_{j \in [1,1/\delta]}\), it is bounded by

\[
\sup_{\theta' \in [\theta - \delta/100, \theta + \delta/100]} \sup_{Z \in \mathbb{R}} \mathbb{P}\left( \left| \sum_{j=1}^{1/\delta} c_j \xi_j r^j \cos(j\theta') - Z \right| \leq \exp\left(-\delta^{-c_1}/2\right) \right) \ll \delta^{2\alpha}.
\]

Applying Lemma 3.6 with \(B = 4\alpha, E = [1, 1/\delta], M = 1/\delta, I = [\theta - \delta/100, \theta + \delta/100], e_j = c_j r^j\) and \(\tilde{e} = \delta\) (where we use Condition (A3) and the assumption that \(r = |z| \geq 1 - 3\delta\) to get \(|e_j| \geq \tilde{e}\)), we obtain \(\theta' \in [\theta - \delta/100, \theta + \delta/100]\) such that for all \(Z \in \mathbb{C}\),

\[
\mathbb{P}\left( \left| \sum_{j=1}^{1/\delta} c_j \xi_j r^j \cos(j\theta') - Z \right| \leq \exp\left(-\delta^{-c_1}/2\right) \right) \ll \delta^{2\alpha}.
\]

This proves that (2) holds with probability at least 1 \(-\frac{\delta^{-4} - s_0}{m^2 m^{-4}}\), concluding the proof of (96). \(\square\)

Proof of Lemma 8.7. Since \(x_1 \in B(x, \delta/100)\), it suffices to show that

\[
\int_{B(x_1, 2\delta/3 + \delta/100)} (\log |P_n(z)|)^2 \, dz \ll \delta^{-8c_1 + 2}.
\]

By (95), there exists an \(r \in [2\delta/3 + \delta/100, 3\delta/4 - \delta/100]\) such that \(P_n\) does not have zeros in the (closed) annulus \(B(x_1, r + \eta) \setminus B(x_1, r - \eta)\) with center at \(x_1\) and radii \(r \pm \eta\) where \(\eta \gg \delta^{1 + c_1}\).

It’s now sufficient to show that

\[
\int_{B(x_1, r)} (\log |P_n(z)|)^2 \, dz \ll \delta^{-8c_1 + 2}.
\]
Let $\zeta_1, \ldots, \zeta_k$ be all zeros of $P_n$ in $B(x_1, r - \eta)$, then $k \ll \delta^{-c_1}$ and $P_n(z) = (z - \zeta_1) \ldots (z - \zeta_k)g(z)$ where $g$ is a polynomial having no zeros on the closed ball $B(x_1, r + \eta)$. By triangle inequality,

$$
\left( \int_{B(x_1, r)} \log^2 |P_n(z)| \, dz \right)^{1/2} \leq \sum_{i=1}^{k} \left( \int_{B(x_1, r)} \log^2 |z - \zeta_i| \, dz \right)^{1/2} + \left( \int_{B(x_1, r)} \log^2 |g(z)| \, dz \right)^{1/2}
$$

(101)

$$
\ll \delta^{1-2c_1} + \left( \int_{B(x_1, r)} \log^2 |g(z)| \, dz \right)^{1/2},
$$

where in the last inequality, we used

$$
\int_{B(x_1, r)} \log^2 |z - \zeta_i| \, dz \leq \int_{B(0, 3\delta/2)} \log^2 |z| \, dz \ll \delta^{2-2c_1}.
$$

Next, we will bound $\int_{B(x_1, r)} \log^2 |g(z)| \, dz$ by finding a uniform upper bound and lower bound for $\log |g(z)|$. Since $\log |g(z)|$ is harmonic in $B(x_1, r)$, it attains its extrema on the boundary. Thus,

$$
\left( \int_{B(x_1, r)} \log^2 |g(z)| \, dz \right)^{1/2} \ll \delta \max_{z \in \partial B(x_1, r)} |\log |g(z)||.
$$

(102)

Notice that $\log |g(z)|$ is also harmonic on the ball $B(x_1, r + \eta)$. For the upper bound of $\log |g(z)|$, we prove the following.

**Lemma 8.3.** For every $z$ in $B(x_1, r + \eta)$, we have

$$
\log |g(z)| \leq \delta^{-2c_1}.
$$

**Proof.** Since a harmonic function attains its extrema on the boundary, we can assume that $z \in \partial B(x_1, r + \eta)$. By Condition (T1), $\log |P_n(z)| \leq \delta^{-c_1}$. Additionally, by noticing that $|z - \zeta_i| \geq 2\eta$ for all $1 \leq i \leq k$, we get

$$
\log |g(z)| = \log |P_n(z)| - \sum_{i=1}^{k} \log |z - \zeta_i| \leq \delta^{-c_1} - k \log(2\eta) \leq \delta^{-2c_1}
$$

(103)

as desired. \qed

As for the lower bound, let $u(z) = \delta^{-2c_1} - \log |g(z)|$, then $u$ is a non-negative harmonic function on the ball $B(x_1, r + \eta)$. By Harnack’s inequality (see [32, Chapter 11]) for the subset $B(x_1, r)$ of the above ball, we have that for every $z \in B(x_1, r)$,

$$
\alpha u(x_1) \leq u(z) \leq \frac{1}{\alpha} u(x_1),
$$

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

$$
\alpha \left( \delta^{-2c_1} - \log |g(x_1)| \right) \leq \delta^{-2c_1} - \log |g(z)| \leq \frac{1}{\alpha} \left( \delta^{-2c_1} - \log |g(x_1)| \right).
$$

And so,

$$
|\log |g(z)|| \leq \frac{1}{\alpha} |\log |g(x_1)|| + \frac{1}{\alpha} \delta^{-2c_1} \ll \delta^{-c_1} |\log |g(x_1)|| + \delta^{-3c_1}.
$$

(104)
Thus, we reduce to bounding $|\log |g(x)||$. From Lemma 8.3 and condition (12), we have
\[
\delta^{-2c_1} \geq \log |g(x)| = \log |P(x)| - \sum_{i=1}^{k} \log |x_{1} - \zeta_{i}| \geq \log |P(x)| \geq -\frac{1}{2} \delta^{-c_1}.
\]
And so, $|\log |g(x)|| \leq \delta^{-2c_1}$, which together with (104) give
\[
|\log |g(z)|| \ll \delta^{-3c_1}.
\]

From (101), (102), and (105), we obtain (100) and hence Lemma 8.1. \hfill \Box

### 8.2. Proof of (21)
We first reduce to the hyperbolic polynomials for which the Kac-Rice formula (20) is easier to handle. Consider the hyperbolic polynomial with coefficients $c_{j,\rho} := \sqrt{\frac{(2p+1)\cdots(2p+j)}{j!}}$, $0 \leq j \leq n$.

By condition (1), $c_{j,\rho} = \Theta(c_{j})$ for all $j \geq N_0$. Using the Kac-Rice formula (20), we have
\[
\mathbb{E} \left( \tilde{N}_n(S_2 \cup S_3) \right) \ll \int_{S_2 \cup S_3} \sqrt{\sum_{j=0}^{n} \sum_{k=j+1}^{n} c_{j,\rho}^2 c_{k,\rho}^2 (k-j)^{2t_{2j}+2k-2}} \sum_{j=0}^{n} c_{j,\rho}^2 t_{2j} \, dt.
\]

We use [6, Lemma 10.3] with $h(k) = c_{k,\rho}^2$ which estimates the above integrand uniformly over the interval $(1 - \frac{1}{C}, 1 - \frac{C}{n})$ for some sufficiently large constant $C$ and asserts that
\[
\sqrt{\sum_{j=0}^{n} \sum_{k=j+1}^{n} c_{j,\rho}^2 c_{k,\rho}^2 (k-j)^{2t_{2j}+2k-2}} \ll \frac{\sqrt{2p+1}}{2\pi(1-t)} + (1-t)^{\rho-1/2} + \frac{1}{n(1-t)^2} \ll \frac{1}{1-t}.
\]

This together with (106) give (21) for $\delta_i \geq \frac{2C}{n}$ as in this case, $S_2 \cup S_3 \subset (1 - \frac{1}{C}, 1 - \frac{C}{n})$.

If $\delta_i \leq \frac{2C}{n}$, since $k-j \leq n$, for all $t \in S_2 \cup S_3$, we have
\[
\sqrt{\sum_{j=0}^{n} \sum_{k=j+1}^{n} c_{j,\rho}^2 c_{k,\rho}^2 (k-j)^{2t_{2j}+2k-2}} \ll n.
\]

Plugging this into (106) and using the fact that $2C/n \geq \delta_i \geq \delta_M \geq 1/n$ give
\[
\mathbb{E} \left( \tilde{N}_n(S_2 \cup S_3) \right) \ll n\delta_i^{1+\alpha} \ll n^{-\alpha} \ll \delta_i^\alpha
\]
and hence (21) for $\delta_i \leq \frac{2C}{n}$, completing the proof of (21) for all values of $\delta_i$. \hfill \Box

### 8.3. Proof of Lemma 3.3
In this section, we deduce Lemma 3.3 from Lemma 3.7

The constant $\alpha_0$ in this proof will be a small fraction of the $\alpha_0$ in Lemma 3.7. Let $\tilde{K}(x_{ik})_{ik} := K(x_{ik} + \frac{1}{2} \log V(w_{ik}))_{ik}$. Then, $\tilde{K}$ still satisfies (16) and we can reduce the problem to showing that
\[
|\mathbb{E} \tilde{K} \left( \log \frac{|P(w_{ik})|}{\sqrt{V(w_{ik})}} \right)_{ik} - \mathbb{E} \tilde{K} \left( \log \frac{|\tilde{P}(w_{ik})|}{\sqrt{V(w_{ik})}} \right)_{ik}| = O(\delta_0^\alpha).
\]

Ideally, we would like to set $L(z_{ik})_{ik} := \tilde{K}(\log |z_{ik}|)_{ik}$ and apply Lemma 3.7 for this function $L$. However, the singularity of the log function at 0 prevents $L$ from satisfying (16). To handle this difficulty, we split the space of $(\log |z_{ik}|)_{ik}$ into two regions $\Omega_1$ and $\Omega_2$ where $\Omega_1$ is the image of
the log function around 0 and show that the contribution from $\Omega_1$ is insignificant. On $\Omega_2$, the log function is well-behaved and we can then apply Lemma 3.7 there.

More specifically, for $M_i := \log (\delta_i^{-12\alpha})$, let

$$\Omega_1 = \{(x_{ik})_{ik} \in \mathbb{R}^{m_1+\cdots+m_M} : x_{ik} \leq -M_i \text{ for some } i, k\}$$

and

$$\Omega_2 = \{(x_{ik})_{ik} \in \mathbb{R}^{m_1+\cdots+m_M} : x_{ik} \geq -M_i - 1 \text{ for all } i, k\}.$$

Let $\psi : \mathbb{R}^{m_1+\cdots+m_M} \to [0, 1]$ be a smooth function taking values in $[0, 1]$ such that $\psi$ is supported in $\Omega_2$, $\psi = 1$ on the complement of $\Omega_1$ and $\|\partial^a \psi\|_{\infty} = O(1)$ for all $0 \leq a \leq 3$. Put $\phi := 1 - \psi$, $K_1 := \bar{K}.\phi$, and $K_2 := \bar{K}.\psi$. We have $\bar{K} = K_1 + K_2$ and both $K_1, K_2$ satisfy [16] with supp $K_1 \subset \Omega_1$, supp $K_2 \subset \Omega_2$.

We now show that the contribution from $K_1$ is negligible. Set $\bar{K}_1 := \|\bar{K}\|_{1,\infty}$ and

$$L_1(z_{ik})_{ik} := \bar{K}_1(\log |z_{ik}|)_{ik}.$$

Since $\|K_1\|_{\infty} \leq \|\bar{K}\|_{1,\infty} \ll 1$, we observe that $L_1$ satisfies

- $|K_1(\log |z_{ik}|)_{ik}| \leq L_1(z_{ik})_{ik}$,
- $\text{supp}(L_1) \subset \{(z_{ik})_{ik} \in \mathbb{C}^{m_1+\cdots+m_M} : |z_{ik}| \leq e^{-M_i} \text{ for some } i, k\}$,
- $L_1$ is constant on $\{(z_{ik})_{ik} \in \mathbb{C}^{m_1+\cdots+m_M} : |z_{ik}| \leq e^{-M_i-1} \text{ for some } i, k\}$,
- $L_1$ satisfies [16] (with the power $2\alpha$ being replaced by $14\alpha$ but that doesn’t affect the argument).

Choose $\alpha_0$ to be small enough such that $C\alpha_0$ is at most the constant $\alpha_0$ in Lemma 3.7 where $C$ is some sufficiently large absolute constant. Applying Lemma 3.7, we get

$$\mathbb{E}\left[K_1 \left(\log \frac{|P(w_{ik})|}{\sqrt{V(w_{ik})}}\right)_{ik}\right] \leq \mathbb{E}L_1 \left(\frac{P(w_{ik})}{\sqrt{V(w_{ik})}}\right)_{ik} \leq \mathbb{E}L_1 \left(\frac{\tilde{P}(w_{ik})}{\sqrt{V(w_{ik})}}\right)_{ik} + O(\delta_0^{\alpha_0}).$$

Since the variables $\tilde{\xi}_i$ are Gaussian, we have

$$\mathbb{E}L_1 \left(\frac{\tilde{P}(w_{ik})}{\sqrt{V(w_{ik})}}\right)_{ik} \ll \mathbb{P}\left(\exists i_k : \left|\tilde{P}(w_{ik})\right| \leq e^{-M_i} \leq e^{-M_i} \ll e^{-M_i}\right) \ll \sum_{i=1}^M m_i \delta_i^{12\alpha} \ll \delta_0^{\alpha_0}.$$

Thus, $\mathbb{E}\left[K_1 \left(\log \frac{|P(w_{ik})|}{\sqrt{V(w_{ik})}}\right)_{ik}\right] \ll \delta_0^{\alpha_0}$. Finally, we will show that

$$\mathbb{E}K_2 \left(\log \frac{|P(z_1)|}{\sqrt{V(z_1)}}, \ldots, \log \frac{|P(z_m)|}{\sqrt{V(z_m)}}\right) - \mathbb{E}K_2 \left(\log \frac{\tilde{P}(z_1)}{\sqrt{V(z_1)}}, \ldots, \log \frac{\tilde{P}(z_m)}{\sqrt{V(z_m)}}\right) \ll \delta_0^{\alpha_0}.$$

Define $L_2 : \mathbb{C}^{m_1+\cdots+m_M} \to \mathbb{R}$ by $L_2(z_{ik}) = K_2(\log |z_{ik}|)$. Since supp $K_2 \subset \Omega_2$,

$$\text{supp } L_2 \subset \{(z_{ik})_{ik} : |z_{ik}| \geq e^{-M_i-1} \gg \delta_i^{12\alpha} \text{ for all } i, k\}. $$
Thus, $L_2$ is well-defined and satisfies (16) (with the power $2\alpha$ being replaced by $14\alpha$). Applying Lemma 3.7 gives

$$
\mathbb{E}K_2 \left( \log \frac{|P(w_{ik})|}{\sqrt{V(w_{ik})}} \right)_{ik} - \mathbb{E}K_2 \left( \log \frac{|\tilde{P}(w_{ik})|}{\sqrt{V(w_{ik})}} \right)_{ik} = \mathbb{E}L_2 \left( \frac{|P(w_{ik})|}{\sqrt{V(w_{ik})}} \right)_{ik} - \mathbb{E}L_2 \left( \frac{|\tilde{P}(w_{ik})|}{\sqrt{V(w_{ik})}} \right)_{ik} \ll \delta_0^\alpha.
$$

This completes the proof of Lemma 3.3. □

8.4. Proof of (56). In this section, we prove (56), namely, for a sufficiently large constant $C$, we have

$$
(108) \quad V(x) := \sum_{i=0}^{n} c_i^2 x^{2i} = \frac{\Theta(1)}{(1-x+1/n)^{2\rho+1}} \quad \forall x \in (1-1/C, 1).
$$

To this end, we will repeatedly use (A3) and the assumption that $\rho > -1/2$.

If $x \geq 1 - \frac{1}{n}$, we have

$$
V(x) \leq \sum_{i=0}^{n} i^2 c_i^2 \ll \sum_{i=0}^{N_0} i + \sum_{i=N_0}^{n} i^{2\rho} \ll n^{2\rho+1}.
$$

For the lower bound, we have $x^{2i} \geq (1 - \frac{1}{n})^{2n} \gg 1$ and so

$$
V(x) \gg \sum_{i=0}^{n} c_i^2 \gg \sum_{i=N_0}^{n} i^{2\rho} \gg n^{2\rho+1}.
$$

These bounds prove (108) for $x \geq 1 - \frac{1}{n}$.

If $1 - \frac{1}{n} < x < 1 - \frac{1}{C}$, letting $L = \frac{1}{1-x} \in (C, n)$, we have $\frac{1}{(1-x+1/n)^{2\rho+1}} = \Theta(L^{2\rho+1})$ and

$$
V(x) \gg \sum_{i=L}^{2L} c_i^2 x^{2i} \gg \sum_{i=L}^{2L} i^{2\rho} \gg L^{2\rho+1}.
$$

As for the upper bound, we have for any constant $C'$,

$$
(109) \quad V(x) \ll \sum_{i=0}^{N_0} 1 + \sum_{i=N_0}^{\infty} i^{2\rho} x^{2i} \ll 1 + \sum_{i=C'L}^{\infty} i^{2\rho} x^{2i} + \sum_{i=C'L}^{\infty} i^{2\rho} x^{2i} \ll L^{2\rho+1} + \sum_{i=C'L}^{\infty} i^{2\rho} x^{2i}.
$$

Since $x = 1 - \frac{1}{L} \leq e^{-1/L}$, the right-most sum is at most

$$
(110) \quad \sum_{i=C'L}^{\infty} i^{2\rho} x^{2i} \leq \sum_{i=C'L}^{\infty} i^{2\rho} e^{-2i/L} = L^{2\rho} \sum_{i=C'L}^{\infty} \left( \frac{i}{L} \right)^{2\rho} e^{-2i/L}.
$$

By choosing $C'$ sufficiently large (depending only on $\rho$) such that the function $t \to t^{2\rho} e^{-2t}$ is decreasing on $(C' - 1, \infty)$, we have

$$
(111) \quad \sum_{i=C'L}^{\infty} \left( \frac{i}{L} \right)^{2\rho} e^{-2i/L} \leq \int_{C'L-1}^{\infty} \left( \frac{s}{L} \right)^{2\rho} e^{-2s/L} ds = L \int_{C'-1/L}^{\infty} t^{2\rho} e^{-2t} dt \ll L.
$$

Plugging this into (109) and (110), we obtain $V(x) \ll L^{2\rho+1}$ which is the desired upper bound. □
8.5. Proof of (73) and (74).

Proof of (73). We need to show that

$$\sum_{0 \leq i < k \leq n} c_i^2 c_k^2 (x^i y^k - x^k y^i)^2 = \Theta \left( \sum_{0 \leq i < k \leq n} c_i^2 c_k^2 (x^i y^k - x^k y^i)^2 \right).$$

By Condition (A3), $c_i \ll c_{i,\rho}$ for all $i \geq 0$, and so the left-hand side of (112) is at most the order of the right-hand side. To prove the reverse, by Condition (A3), $c_i \gg c_{i,\rho}$ for all $i \geq N_0$, so

$$\sum_{0 \leq i < k \leq n} c_i^2 c_k^2 (x^i y^k - x^k y^i)^2 \leq \sum_{N_0 \leq i < k \leq n} c_i^2 c_k^2 (x^i y^k - x^k y^i)^2 \gg \sum_{N_0 \leq i < k \leq n} c_i^2 c_k^2 (x^i y^k - x^k y^i)^2.$$

Thus, it remains to show that the remaining terms on the right-hand side of (112) are of smaller order, namely, for all $1 \leq i < N_0$,

$$\sum_{k=0}^{n} c_{i,\rho}^2 c_{k,\rho}^2 (x^i y^k - x^k y^i)^2 \ll \sum_{N_0 \leq j \leq n} c_j^2 c_{k,\rho}^2 (x^j y^k - x^k y^j)^2.$$

Since $N_0$ is a constant, $c_{k,\rho} = \Theta(c_{k+N_0,\rho})$ for all $k \geq 0$ and since $x y = \Theta(1)$, we have for $j' = i + N_0$,

$$\sum_{k=0}^{n} c_{i,\rho}^2 c_{k,\rho}^2 (x^i y^k - x^k y^i)^2 \ll \sum_{k=0}^{n} c_{j',\rho}^2 c_{k+N_0,\rho}^2 (x^{j'} y^{k+N_0} - x^k y^{j'})^2 = \sum_{k=0}^{n} c_{j',\rho}^2 c_{k,\rho}^2 (x^{j'} y^k - x^k y^{j'})^2.$$

Assume without loss of generality that $x < y$. Using the simple observation that

$$0 \leq y^{j+1} - x^{j+1} \leq 2 (y^j - x^j) \forall j \geq 1,$$

we have

$$\sum_{k=n+1}^{n+N_0} c_{j',\rho}^2 c_{k,\rho}^2 (x^{j'} y^k - x^k y^{j'})^2 \ll \sum_{k=n+1-N_0}^{n} c_{j',\rho}^2 c_{k,\rho}^2 (x^{j'} y^k - x^k y^{j'})^2.$$

And so, the right-most side of (114) is of order at most the right-most side of (113), proving (112).

Proof of (74). We want to show that for every $x \in [1 - a_n, 1 - b_n]$,

$$\sum_{k=0}^{n} c_{k,\rho}^2 x^{2k} = \frac{1 + O(\varepsilon_0)}{(1 - x^2)^{2\rho+1}},$$

where $\varepsilon_0 = \exp (- (\log \log n)^{1+2\varepsilon})$. By Taylor’s expansion, we have

$$S := \sum_{k=0}^{\infty} c_{k,\rho}^2 x^{2k} = \frac{1}{(1 - x^2)^{2\rho+1}}.$$

Thus, it suffices to show that

$$\sum_{k=n+1}^{\infty} c_{k,\rho}^2 x^{2k} \ll \varepsilon_0 S.$$

We have

$$\sum_{k=n+1}^{\infty} c_{k,\rho}^2 x^{2k} = x^{2n+2} \sum_{k=0}^{\infty} c_{n+1+k,\rho}^2 c_{k,\rho}^2 x^{2k}.$$
and so, it is left to verify that for all \(k \geq 0\),

\[
x^{2n} \frac{c_{n+1+k,\rho}}{c_{k,\rho}} \ll \varepsilon_0.
\]

Indeed, we have

\[
x^{2n} \sum_{i=1}^{n+1} \frac{2\rho + k + i}{k + i} \leq x^{2n} \prod_{i=1}^{n+1} \frac{2\rho + i}{i} \leq x^{2n} \prod_{i=1}^{n+1} \frac{2[\rho] + i + 1}{i} = x^{2n} (n + 2 \ldots (n + 2 + 2[\rho])/(2[\rho] + 1)! \ll x^{2n} n^{2\rho+1}.
\]

Using \(x \leq 1 - b_n \leq 1 - (\log n)^2/n\) by the assumption (6), we obtain

\[
x^{2n} \frac{c_{n+1+k,\rho}}{c_{k,\rho}} \ll \left(1 - \frac{(\log n)^2}{n}\right)^{2n} n^{2\rho+1} \ll \exp(-2(\log n)^2) n^{2\rho+1} \ll \varepsilon_0.
\]

\[\square\]

8.6. Proof of Lemma 6.2 Since \(b_n \geq 1/n\), for all \(x \in [1 - a_n, 1 - b_n], 1 - x \geq b_n \geq 1/n\). We write \(x = 1 - \frac{1}{L}\).

By (56), on the right-most side of (62), we have

\[
\text{Var} \ P_n(x) = \frac{\Theta(1)}{(1 - x)^{2\rho+1}} = \Theta(L^{2\rho+1}).
\]

On the other side, we have

\[
\text{Var} \ P_n(x) - \text{Var} \ Q_n(x) \leq \sum_{i=0}^{N_0} c_i^2 x^{2i} + \sum_{i=M_x}^{m_x} c_i^2 x^{2i} + \sum_{i=M_x}^{n} c_i^2 x^{2i} \ll 1 + \sum_{i=0}^{M_x} i^{2\rho} + \sum_{i=M_x}^{n} i^{2\rho} e^{-2i/L}.
\]

By the same argument as in (111), the right-most sum is at most

\[
\sum_{i=M_x}^{n} i^{2\rho} e^{-2i/L} \ll L^{2\rho+1} \int_{M_x/L-1}^{\infty} t^{2\rho} e^{-2t} dt \ll L^{2\rho+1} e^{-\alpha \log \log L} \ll (\log L)^{-\alpha} \mathbb{E}P_n^2(x) \ll A^{-\alpha \rho'} \mathbb{E}P_n^2(x)
\]

where we used \(M_x = \alpha L \log \log L\) and \(A_x = \log L\) by the definition of \(M_x\) and \(A_x\). Thus,

\[
\text{Var} \ P_n(x) - \text{Var} \ Q_n(x) \ll 1 + m_x^{2\rho+1} + A_x^{-\alpha \rho'} \mathbb{E}P_n^2(x) \ll A_x^{-\alpha \rho'} \mathbb{E}P_n^2(x)
\]

where we used \(m_x = L(\log L)^{-\alpha}\) by the definition of \(m_x\). This proves (62).

As for the second part of Lemma 6.2 writing \(x = 1 - 1/L\) and \(y = 1 - 1/K\), we have \(1 \ll L \leq K \leq n\) and \(\log \frac{K}{L} \geq 2\alpha \log \log n\), so

\[M_x = \alpha L \log \log L \leq L \log^\alpha n \leq K \log^{-\alpha} n \leq K \log^{-\alpha} K = m_y.
\]

\[\square\]
8.7. Proof of \((82)\). Let \((c, d) := (a_j, a_{j+1})\). We want to show that for any interval \((c, d) \subset [1 - 1/C, 1]\) with \(d - c \leq D^{-1/2}\),

\[
\mathbb{P} \left( \int_c^d \int_c^d |P''_n(u)| du \geq D^{p-1} \right) \ll D^{-1}.
\]

Let \(I\) denote the above double integral. By Markov’s inequality and Hölder’s inequality, for a large constant \(h\) to be chosen, we have

\[
\mathbb{P}(I \geq D^{p-1}) \leq D^{(1-p)h} \mathbb{E}^I h \leq D^{(1-p)h} (d - c)^{2(h-1)} \mathbb{E} \int_c^d \int_c^d |P''_n(u)| du \geq D^{-p} \max_{u \in [c, d]} \mathbb{E}|P''_n(u)|^h.
\]

and so,

\[
\mathbb{P}(I \geq D^{p-1}) \ll D^{(1-p)h}(d - c)^{2h} \max_{u \in [c, d]} \mathbb{E}|P''_n(u)|^h \leq D^{-ph} \max_{u \in [c, d]} \mathbb{E}|P''_n(u)|^h.
\]

Since \(P''_n(u)\) is a Gaussian random variable, by the hypercontractivity of Gaussian distribution, we have

\[
\mathbb{E}|P''_n(u)|^h \ll \left( \mathbb{E}|P''_n(u)|^2 \right)^{\frac{2+h}{2}} \ll n^{O(1)}.
\]

Thus, by choosing \(h = 2/p\),

\[
\mathbb{P}(I \geq D^{p-1}) \ll D^{-ph} n^{O(1)} \ll D^{-1}
\]

where we used \(D = \exp \left( a_n^{-1}/100 \right) \gg n^n\) for any constant \(C\) as \(a_n\) satisfies Condition \((2)\). \(\square\)

8.8. Proof of \((85)\). We have

\[
\text{Cov} \left( N^{(2)}_{P_n} (a, b), N^{(2)}_{R_n} (c, d) \right) = \mathbb{E}N^{(2)}_{P_n} (a, b) N^{(2)}_{R_n} (c, d) - \mathbb{E}N^{(2)}_{P_n} (a, b) \cdot \mathbb{E}N^{(2)}_{R_n} (c, d).
\]

Thus, we have by definition of \(N^{(2)}\) that

\[
\text{Cov} \left( N^{(2)}_{P_n} (a, b), N^{(2)}_{R_n} (c, d) \right) = \frac{1}{4\pi^4} \int_a^b \int_c^d \int_D^2 \int_D^2 \int_D^D \int_D^D \frac{1}{uv^2} \cdot \left[ \mathbb{E} F_1 (x, u, s) F_2 (y, v, t) - \mathbb{E} F_1 (x, u, s) \mathbb{E} F_2 (y, v, t) \right] dt ds dv du dx
\]

where

\[
F_1 (x, u, s) := \left( 1 - \cos \left( uP_n'(x) \right) \right) \cos (sP_n(x)), \quad F_2 (y, v, t) := \left( 1 - \cos \left( vR_n'(y) \right) \right) \cos (tR_n(y)).
\]

Note that we can use Fubini’s theorem in the above calculation because the integrands are absolutely integrable.

We have

\[
F_1 (x, u, s) F_2 (y, v, t) = \left( 1 - \cos \left( uP_n'(x) \right) \right) \cos (sP_n(x)) \left( 1 - \cos \left( vR_n'(y) \right) \right) \cos (tR_n(y))
\]

\[
= \cos (sP_n(x)) \cos (tR_n(y)) - \cos \left( uP_n'(x) \right) \cos (sP_n(x)) \cos (tR_n(y))
\]

\[
- \cos \left( vR_n'(y) \right) \cos (sP_n(x)) \cos (tR_n(y)) + \cos \left( uP_n'(x) \right) \cos (vR_n'(y)) \cos (sP_n(x)) \cos (tR_n(y))
\]

\[
= \frac{1}{2} \sum \cos (sP_n(x) + tR_n(y)) - \frac{1}{4} \sum \cos (uP_n'(x) + sP_n(x) + tR_n(y))
\]

\[
- \frac{1}{4} \sum \cos (vR_n'(y) + sP_n(x) + tR_n(y)) + \frac{1}{8} \sum \cos (uP_n'(x) + vR_n'(y) + sP_n(x) + tR_n(y))
\]

We recall that the random variables \(\xi_i\) are iid standard Gaussian and for a standard Gaussian random variable \(Z\) and any real number \(a\), \(\mathbb{E} \cos (aZ) = E e^{i a Z} = e^{-a^2/2} = \exp \left( -\frac{1}{2} \text{Var} \ (aZ) \right)\).
Thus,
\[
\mathbb{E} \mathcal{F}_1(x, u, s) \mathcal{F}_2(y, v, t) = \frac{1}{2} \sum \exp \left( -\frac{1}{2} \text{Var} \ (s P_n(x) \pm t R_n(y)) \right) \\
- \frac{1}{4} \sum \exp \left( -\frac{1}{2} \text{Var} \ (u P'_n(x) \pm s P_n(x) \pm t R_n(y)) \right) \\
- \frac{1}{4} \sum \exp \left( -\frac{1}{2} \text{Var} \ (v R'_n(y) \pm s P_n(x) \pm t R_n(y)) \right) \\
+ \frac{1}{8} \sum \exp \left( -\frac{1}{2} \text{Var} \ (u P'_n(x) \pm v R'_n(y) \pm s P_n(x) \pm t R_n(y)) \right).
\]

Similarly,
\[
\mathbb{E} \mathcal{F}_1(x, u, s) = \exp \left( -\frac{1}{2} \text{Var} \ (s P_n(x)) \right) - \frac{1}{2} \sum \exp \left( -\frac{1}{2} \text{Var} \ (u P'_n(x) \pm s P_n(x)) \right)
\]
and
\[
\mathbb{E} \mathcal{F}_2(y, v, t) = \exp \left( -\frac{1}{2} \text{Var} \ (t R_n(y)) \right) - \frac{1}{2} \sum \exp \left( -\frac{1}{2} \text{Var} \ (v R'_n(y) \pm t R_n(y)) \right).
\]

Plugging these formulas into (117), we obtain (85). \[\square\]

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

E-mail address: onguyen@princeton.edu

Department of Mathematics, Yale University, New Haven, CT 06520, USA

E-mail address: van.vu@yale.edu