



PROJECT MUSE®

Central limit theorems for the real zeros of Weyl polynomials

Yen Do, Van Vu

American Journal of Mathematics, Volume 142, Number 5, October
2020, pp. 1327-1369 (Article)

Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/ajm.2020.0034>

AMERICAN JOURNAL
OF MATHEMATICS

EDITED BY
DAVID B. FRENKEL
AND
DAVID P. ROSENBLUTH

BOARD OF EDITORS
DAVID B. FRENKEL
DAVID P. ROSENBLUTH
DAVID B. FRENKEL
DAVID P. ROSENBLUTH

PRODUCTION AND CIRCULATION
BY
JOHNS HOPKINS UNIVERSITY PRESS

VOLUME 142, NUMBER 5
OCTOBER 2020

JOHNS HOPKINS
UNIVERSITY PRESS

➔ *For additional information about this article*

<https://muse.jhu.edu/article/763847/summary>

CENTRAL LIMIT THEOREMS FOR THE REAL ZEROS OF WEYL POLYNOMIALS

By YEN DO and VAN VU

Abstract. We establish the central limit theorem for the number of real roots of the Weyl polynomial $P_n(x) = \xi_0 + \xi_1 x + \cdots + \frac{1}{\sqrt{n}!} \xi_n x^n$, where ξ_i are iid Gaussian random variables. The main ingredients in the proof are new estimates for the correlation functions of the real roots of P_n and a comparison argument exploiting local laws and repulsion properties of these real roots.

1. Introduction. In this paper, we discuss random polynomials with Gaussian coefficients, namely, polynomials of the form

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

where ξ_i are iid standard normal random variables, and c_i are real, deterministic coefficients (which can depend on both i and n).

The central object in the theory of random polynomials, starting with the classical works of Littlewood and Offord [16, 17, 18], is the distribution of the real roots. This will be the focus of our paper. In what follows, we denote by N_n the number of real roots of P_n .

One important case is when $c_1 = \cdots = c_n = 1$. In this case, the polynomial is often referred to as Kac polynomial. Littlewood and Offord [16, 17, 18] in the early 1940s, to the surprise of many mathematicians of their time, showed that N_n is typically polylogarithmic in n .

THEOREM 1. (Littlewood-Offord) *For Kac polynomials,*

$$\frac{\log n}{\log \log n} \leq N_n \leq \log^2 n$$

with probability $1 - o(1)$.

Almost simultaneously, Kac [15] discovered his famous formula for the density function $\rho(t)$ of N_n ; he show

Manuscript received August 1, 2017; revised May 15, 2018.

Research of the first author supported in part by NSF grant DMS-1521293; research of the second author supported in part by NSF grant DMS-0901216.

American Journal of Mathematics 142 (2020), 1327–1369. © 2020 by Johns Hopkins University Press.

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

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

Consequently,

$$(1.2) \quad \mathbb{E}N_n = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} |y| p(t, 0, y) dy.$$

For Kac polynomials, he computed $p(t, 0, y)$ explicitly and showed [15]

$$(1.3) \quad \mathbb{E}N_n = \frac{1}{\pi} \int_{-\infty}^{\infty} \sqrt{\frac{1}{(t^2 - 1)^2} + \frac{(n+1)^2 t^{2n}}{(t^{2n+2} - 1)^2}} dt = \left(\frac{2}{\pi} + o(1) \right) \log n.$$

More elaborate analysis of Wilkins [30] and also Edelman and Kostlan [8] provide a precise estimate of the RHS, showing

$$(1.4) \quad \mathbb{E}N_n = \frac{2}{\pi} \log n + C + o(1),$$

where $C = 0.65 \dots$ is an explicit constant.

The problem of estimating the variance and establishing the limiting law has turned out to be significantly harder. Almost thirty years after Kac's work, Maslova solved this problem.

THEOREM 2. [19, 20] *Consider Kac polynomials. We have, as n tends to infinity*

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

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

Both Kac's and Maslova's results hold in a more general setting where the gaussian variable is replaced by any random variable with the same mean and variance; see [13, 19, 20].

Beyond the case $c_1 = \dots = c_n = 1$, the expectation of N_n is known for many other settings, see for instance [4, 8, 9, 19] and the references therein and also the introduction of [7] for a recent update. In many cases, the order of magnitude of the coefficients c_i (rather than their precise values) already determines the expectation $\mathbb{E}N_n$ almost precisely (see the introduction of [7]).

The limiting law is a more challenging problem, and progress has been made only very recently, almost 40 years after the publication of Maslova's result. In 2015, Dalmao [5] established the CLT for Kostlan-Shub-Smale polynomials (the

case when $c_i = \sqrt{\binom{n}{i}}$. It is well known that in this case the expectation $\mathbb{E}N_n$ is precisely $2\sqrt{n}$ [8].

THEOREM 3. [5] *Consider Kostlan-Shub-Smale polynomials. We have, as n tends to infinity*

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

Furthermore $\text{Var } N_n = (K + o(1))\sqrt{n}$, where $K = 0.57\dots$ is an explicit constant.

There are also many recent results on random trigonometric polynomial; see [2, 3, 11, 27]; in fact, [5] is closely related to [2], and the proof of Theorem 3 used the ideas developed for random trigonometric polynomials from [2]. In particular, the papers mentioned above made essential use of properties of gaussian processes.

In this paper, we first establish the central limit theorem for N_n for another important class of random polynomials, the Weyl polynomials

$$P_n(x) = \sum_{k=0}^n \frac{\xi_k}{\sqrt{k!}} x^k.$$

THEOREM 4. *Consider Weyl polynomials. We have, as $n \rightarrow \infty$,*

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

Furthermore $\text{Var } N_n = (2K + o(1))\sqrt{n}$, where $K = 0.18198\dots$ is an explicit constant.

It is well known that for Weyl polynomials $\mathbb{E}N_n = (\frac{2}{\pi} + o(1))\sqrt{n}$ (see e.g., [29, Theorem 5.3]). We give the exact value of K in the next section.

Our method for proving the CLT is new, and it actually yields a stronger result, which establishes the following CLT for a very general class of linear statistics.

To fix notation, let $h : \mathbb{R} \rightarrow \mathbb{R}$. Given $0 < \alpha \leq 1$, we say that h is α -Hölder continuous on an interval $[a, b]$ if $|h(x) - h(y)| \leq C|x - y|^\alpha$ for any $a \leq x, y \leq b$, and the constant C is uniform over x, y . Below let Z_n denote the (multi)set of the real zeros of P_n .

THEOREM 5. *There is a finite positive constant K such that the following holds.*

Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be bounded, nonzero, and supported on $[-1, 1]$ such that

(i) h has finitely many discontinuities and

(ii) h is Hölder continuous when restricted to each interval in the partition of $[-1, 1]$ using these discontinuities.

Let $(R_n) \rightarrow \infty$ such that $R_n \leq n^{1/2} + o(n^{1/4})$ and let $N_n = \sum_{x \in Z_n} h(x/R_n)$. Then

$$\lim_{n \rightarrow \infty} \frac{\text{Var}[N_n]}{R_n \|h\|_2^2} = K.$$

Furthermore, as $n \rightarrow \infty$ we have the following convergence in distribution:

$$(1.5) \quad \frac{N_n - \mathbb{E}N_n}{(\text{Var } N_n)^{1/2}} \longrightarrow N(0, 1).$$

Taking $h = 1_I$ where I is union of finitely many intervals in $[-1, 1]$, we obtain the following corollary, which establishes the CLT for the number of real roots in unions of intervals with total length tending to infinity.

COROLLARY 1. *There is a finite positive constant K such that the following holds. Let $I \subset [-1, 1]$ be union of finitely many intervals. Let $(R_n) \rightarrow \infty$ such that $R_n \leq n^{1/2} + o(n^{1/4})$ and let N_n be the number of zeros of P_n in $R_n I = \{R_n x, x \in I\}$.*

Then

$$\lim_{n \rightarrow \infty} \frac{\text{Var}[N_n]}{R_n |I|} = K.$$

Furthermore, as $n \rightarrow \infty$ we have the following convergence in distribution:

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

For the special case when I is an interval of the form $[-a, a]$, the above asymptotics for the variance of the number of real roots was obtained in [25].

The assumption $(R_n) \rightarrow \infty$ on the length is fairly optimal in the sense that asymptotic normality is unlikely to hold for bounded sequences (R_n) due to the repulsion between nearby real roots and the fact that there are not many real roots inside a bounded interval (see Lemma 4). A similar result of this type was obtained by Granville and Wigman [11] for random trigonometric polynomials, in the special case where the union I consists of one interval.

2. A sketch of our argument and the outline of the paper. The heart of the matter is Theorem 5. It is well known that most of the real roots of the Weyl polynomial (which we will denote by P_n in the rest of the proof) are inside $[-\sqrt{n}, \sqrt{n}]$; see for instance [8, 29] (see also Lemma 4 of the current paper for a local law for the number of real roots of P_n). Instead of considering N_n , we restrict to the number of real roots inside $[-\sqrt{n}, \sqrt{n}]$. By Theorem 5, this variable satisfies CLT. To conclude the proof of Theorem 4, we will use a tool from [29] to bound the

number of roots outside this interval, and show that this extra factor is negligible with respect to the validity of the CLT.

In order to establish Theorem 5, we first prove a central limit theorem for the random Weyl series

$$P_{\infty}(x) = \sum_{k=0}^{\infty} \frac{\xi_k}{\sqrt{k!}} x^k.$$

Let Z denote the (multi)set of the real zeros of P_{∞} where each element in Z is repeated according to its multiplicity.

For $h : \mathbb{R} \rightarrow \mathbb{R}$ and $R > 0$ let $n(R, h) = \sum_{x \in Z} h(x/R)$.

THEOREM 6. *There is a finite positive constant K such that the following holds.*

Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be nonzero compactly supported and bounded. Then

$$\lim_{R \rightarrow \infty} \frac{\text{Var}[n(R, h)]}{R \|h\|_2^2} = K$$

and as $R \rightarrow \infty$ we have the following convergence in distribution:

$$\frac{n(R, h) - \mathbb{E}n(R, h)}{\sqrt{\text{Var}[n(R, h)]}} \rightarrow N(0, 1).$$

Furthermore, for any $k \geq 1$ it holds that $\mathbb{E}[n(R, h)^k] \leq C_{h,k} R^k$.

The constant K is the same in Theorem 4, Theorem 5, and Theorem 6, and could be computed explicitly:

$$(2.1) \quad K = \frac{1}{\pi} + \int \left(\rho(0, t) - \frac{1}{\pi^2} \right) dt$$

where $\rho(s, t)$ is the two-point correlation function for the real zeros of P_{∞} . In fact, numerical computation of K was done by Schehr and Majumdar [25] using an explicit evaluation of $\rho(s, t)$ (from the Kac-Rice formula), giving $K = 0.18198 \dots$. For the convenience of the reader and to keep the paper self-contained we sketch some details in Appendix C.

We deduce Theorem 5 from Theorem 6 via a comparison argument. Roughly speaking, we try to show that, restricted to certain intervals, there is a bijection between the real roots of the two functions. This argument relies critically on the repulsion properties of the real roots of P_n and P_{∞} (see Section 6).

The rest of the paper is devoted to proving Theorem 6. By extending the polynomial to the full series, we can take advantage of the invariance properties of the root process. The main ingredients of the proof are estimates for the correlation functions of the real zeros of P_{∞} . These correlation function estimates are inspired

by related results for the complex zeros of P_∞ by Nazarov and Sodin [22], and we adapt their approach to the real setting. One of the essential steps in [22] is to use a Jacobian formula (which relates the distribution of the coefficients of a polynomial to the distribution of its complex roots) to estimate the correlation functions of random polynomials with fixed degrees. Such formula is, however, not available for real roots, and to overcome this difficulty we use a general expression for correlation functions of real roots of random polynomials due to Götze, Kaliada, Zaporozhets [10]. This expression turns out to be useful to study correlation of small (real) roots, and to remove the smallness assumption we appeal to various invariant properties of the real roots of P_∞ .

Outline of the paper. In Section 3 we will prove several estimates concerning the repulsion properties of the real zeros of P_n and P_∞ . In Section 4 we will prove some local estimates for the real roots of P_n . In Section 6 we will use these estimates to prove Theorem 5 assuming the validity of Theorem 6.

In Section 7 we summarize the new estimates for the correlation functions for the real zeros of P_∞ , which will be used in Section 8 to prove Theorem 5.

The proof of the correlation function estimates stated in Section 7 will be presented in the remaining sections.

Notational convention. By $A \lesssim B$ we mean that there is a finite positive constant C such that $|A| \leq CB$. By $A \lesssim_{t_1, t_2, \dots} B$ we mean that there is a finite positive constant C that may depend on $t_1, t_2, \dots$ such that $|A| \leq CB$. Sometimes we also omit the subscripts when the dependency is clear from the context.

We also say that an event holds with overwhelming probability if it holds with probability at least $1 - O_C(n^{-C})$ where $C > 0$ is any fixed constant.

For any $I \subset \mathbb{R}$ we will let $N_n(I)$ be the number of real roots of P_n in I .

Acknowledgments. The authors would like to thank Manjunath Krishnapur for useful suggestions and Gregory Schehr for a correction concerning the computation of the explicit constant K .

This work was initiated during our visit at the Vietnam Institute for Advanced Study in Mathematics during 2016. We thank the Institute for the hospitality and support.

3. Real root repulsion. In this section we will prove some repulsion estimates for the real roots of P_n (and P_∞). These estimates will be used to deduce Theorem 5 from Theorem 6.

3.1. Uniform estimates for P_n and P_∞ . We first establish several basic estimates for the derivatives of P_n and P_∞ . For convenience of notation let $P_{>n} = P_\infty - P_n$.

LEMMA 1. Let $I_n = [-n^{1/2} + n^{1/6} \log n, n^{1/2} - n^{1/6} \log n]$.

For any $m \geq 0$ integer and $C > 0$ there is a constant $c = c(m, C) > 0$ such that for any $N > 0$ and $n \geq 1$

$$(3.1) \quad \mathbb{P} \left(\sup_{y \in I_n} |e^{-y^2/2} P_n^{(m)}(y)| > N n^{m/2} \right) \lesssim n e^{-cN^2},$$

$$(3.2) \quad \mathbb{P} \left(\sup_{y \in I_n} |e^{-y^2/2} P_\infty^{(m)}(y)| > N n^{m/2} \right) \lesssim n e^{-cN^2},$$

$$(3.3) \quad \mathbb{P} \left(\sup_{y \in I_n} |e^{-y^2/2} P_{>n}^{(m)}(y)| > N n^{-C} \right) \lesssim n e^{-cN^2}.$$

The implicit constants may depend on C and m .

Proof. Without loss of generality we may assume $N > 1$.

We first show (3.1). For any fixed y , we have

$$\begin{aligned} \text{Var}[e^{-y^2/2} P_n^{(m)}(y)] &= e^{-y^2} \sum_{k=m}^n (k+m)^2 \dots (k+1)^2 \frac{y^{2k}}{(k+m)!} \\ &= e^{-y^2} \sum_{k=0}^{n-m} (k+m) \dots (k+1) \frac{y^{2k}}{k!} \\ &< n^m. \end{aligned}$$

Since $e^{-y^2/2} P_n^{(m)}(y)$ is centered Gaussian, it follows that for each fixed y we have

$$\mathbb{P}(|e^{-y^2/2} P_n^{(m)}(y)| \geq N n^{m/2}) \lesssim e^{-N^2/4}.$$

Let $X = (\xi_0, \dots, \xi_n)$ and let $\|\cdot\|$ denote the ℓ_2 norm on $\mathbb{R}^{n+1}$. By Cauchy-Schwarz, we have the deterministic estimate

$$|e^{-y^2/2} P_n^{(m)}(y)| \leq \|X\| \text{Var}[e^{-y^2/2} P_n^{(m)}(y)] < \|X\| n^{m/2}.$$

Let $\delta \in (0, 1)$ to be chosen later. Divide the interval I_n into $O(n^{1/2}\delta^{-1})$ intervals of length at most δ . Let K be the collection of the midpoints of these intervals, then by an union bound we have

$$\mathbb{P}(\exists y \in K : |e^{-y^2/2} P_n^{(m)}(y)| > N n^{m/2}) \leq n^{1/2} \delta^{-1} e^{-N^2/4}.$$

For any $y' \in I_n$, then there is $y \in K$ such that $|y' - y| \leq \delta$. Now, for any $\epsilon \in (0, 1)$, using the mean value theorem we have

$$\begin{aligned}
 & |e^{-(y+\epsilon)^2/2} P_n^{(m)}(y+\epsilon) - e^{-y^2/2} P_n^{(m)}(y)| \\
 & \leq e^{-(y+\epsilon)^2/2} |P_n^{(m)}(y+\epsilon) - P_n^{(m)}(y)| + |e^{-(y+\epsilon)^2/2} - e^{-y^2/2}| P_n^{(m)}(y)| \\
 & \leq \epsilon e^{-(y+\epsilon)^2/2} \sup_{\alpha \in (y, y+\epsilon)} |P_n^{(m+1)}(\alpha)| + \epsilon(y+\epsilon) e^{-y^2/2} |P_n^{(m)}(y)| \\
 & \leq \|X\| \epsilon [n^{(m+1)/2} + (y+\epsilon) n^{m/2}] \\
 & \lesssim \epsilon n^{(m+1)/2} \|X\|.
 \end{aligned}$$

One could crudely estimate $P(\|X\| > e^{N^2/8}) \leq e^{-N^2/4} \mathbb{E}\|X\|^2 = (1+n)e^{-N^2/4}$. (There are sharper estimates for X which follows the chi-squared distribution, but the above estimate is good enough for our purposes.) Therefore by letting $\delta = Nn^{-1/2}e^{-N^2/8}$ and conditioning on the event $\|X\| \leq e^{N^2/8}$ we obtain

$$\begin{aligned}
 & \mathbb{P}(\exists y \in I_n : |e^{-y^2/2} P_n^{(m)}(y)| > 4Nn^{m/2}) \\
 & \leq nN^{-1}e^{-N^2/8} + (n+1)e^{-N^2/4} \\
 & \lesssim ne^{-N^2/8}.
 \end{aligned}$$

This completes the proof of (3.1).

By the triangle inequality it remains to show (3.3).

We proceed as before. Given any fixed y we have

$$\begin{aligned}
 \text{Var} [e^{-y^2/2} P_{>n}^{(m)}(y)] &= e^{-y^2} \sum_{k=n-m+1}^{\infty} (k+1)^2 \dots (k+m)^2 \frac{y^{2k}}{(k+m)!} \\
 &\lesssim y^{2m} e^{-y^2} \sum_{k=n-m+1}^{\infty} \frac{y^{2(k-m)}}{(k-m)!}.
 \end{aligned}$$

Let $y_0 = n^{1/2} - n^{1/6} \log n$. Then for n large enough (relative to m) we have $n - m + 1 \geq y_0^2 > y^2$, consequently for each $k \geq n - m + 1$ the function $h(y) = 2k \log |y| - y^2$ is increasing over $y \in (0, y_0)$. It follows that

$$\text{Var} [e^{-y^2/2} P_{>n}^{(m)}(y)] \lesssim n^m e^{-y_0^2} \sum_{k=n-2m+1}^{\infty} \frac{y_0^{2k}}{k!}.$$

Since $\sqrt{n} \lesssim y_0 \leq \sqrt{n}$ and $m = (1)$, it follows that

$$\begin{aligned} \text{Var} [e^{-y^2/2} P_{>n}^{(m)}(y)] &\lesssim n^m e^{-y_0^2} \sum_{k=n}^{\infty} \frac{y_0^{2k}}{k!} \\ &\lesssim n^m e^{-y_0^2} \sum_{k=n}^{[(1.01)n]} \frac{y_0^{2k}}{k!} \\ &\lesssim n^{m+1} e^{-y_0^2} \frac{y_0^{2n}}{n!} \end{aligned}$$

(here we used the fact that $y_0^2/k < 1/1.01 < 1$ if $k \geq (1.01)n$, and $y^2/k \leq 1$ for $k \geq n$). Consequently,

$$\begin{aligned} \text{Var} [e^{-y^2/2} P_{>n}^{(m)}(y)] &\lesssim n^{m+1} e^{-y_0^2} \frac{y_0^{2n}}{n!} \\ &\lesssim n^{m+1} e^{-y_0^2} \frac{y_0^{2n}}{(n/e)^{n+1/2}} \\ &= n^{m+1/2} e^{-y_0^2 + 2n \log y_0 - n \log n + n}. \end{aligned}$$

Now, for brevity write $y_0 = \sqrt{n}(1 - \beta)$ where $\beta = n^{-1/3} \log n = o(1)$, then

$$\begin{aligned} &-y_0^2 + 2n \log y_0 - n \log n + n \\ &= -n(1 - 2\beta + \beta^2) + n \log n + 2n \log(1 - \beta) - n \log n + n \\ &= 2n \left[\beta - \frac{\beta^2}{2} + \log(1 - \beta) \right] \\ &= -2n \frac{\beta^3}{3} (1 + O(\beta)) \\ &\leq -n\beta^3/3 = -\log n^3/3 \end{aligned}$$

when n is large. Therefore

$$(3.4) \quad \text{Var} [e^{-y^2/2} P_{>n}^{(m)}(y)] \lesssim n^{m+1/2} e^{-\log n^3/3} \lesssim_{C,m} n^{-C}$$

for any $C > 0$. Therefore for any fixed y such that $|y| \leq \sqrt{n}(1 - \log n/n^{1/3})$ we have

$$\mathbb{P}(|e^{-y^2/2} P_{>n}^{(m)}(y)| > N n^{-C}) \lesssim e^{-N^2/4}.$$

Let $X = (\xi_{n+1}, \xi_{n+2}, \dots, \xi_{3n}, \xi_{3n+1}/2, \dots, \xi_m/2^{m-3n}, \dots)$ and let $\|\cdot\|$ denote the ℓ_2 norm on $\mathbb{R}^{n+1}$. By Cauchy-Schwarz and using $y^2 \leq n$, we have the deterministic estimate

$$\begin{aligned} |e^{-y^2/2} P_{>n}^{(m)}(y)| &\leq \|X\| e^{-y^2/2} \left(\sum_{k=n+1-m}^{3n-m} \frac{(k+1) \dots (k+m) y^{2k}}{k!} \right. \\ &\quad \left. + \sum_{k>3n-m} \frac{4^{k+m-3n} (k+1) \dots (k+m) y^{2k}}{k!} \right)^{1/2} \\ &\lesssim \|X\| e^{-y^2/2} \left(\sum_{k=n+1}^{3n} \frac{k^m y^{2k}}{k!} + O\left(\frac{n^m y^{6n}}{(3n)!}\right) \right)^{1/2} \\ &\lesssim \|X\| n^{m/2}. \end{aligned}$$

Let $\delta \in (0, 1)$ to be chosen later. Divide the interval I_n into $O(n^{1/2}\delta^{-1})$ intervals of length at most δ . Let K be the collection of the midpoints of these intervals, then by a union bound we have

$$\mathbb{P}(\exists y \in K : |e^{-y^2/2} P_{>n}^{(m)}(y)| > N n^{-C}) \leq n^{1/2} \delta^{-1} e^{-N^2/4}.$$

For any $y' \in I_n$, then there is $y \in K$ such that $|y' - y| \leq \delta$. Now, for any $\epsilon \in (0, 1)$, using the mean value theorem we have

$$\begin{aligned} &|e^{-(y+\epsilon)^2/2} P_{>n}^{(m)}(y+\epsilon) - e^{-y^2/2} P_{>n}^{(m)}(y)| \\ &\leq e^{-(y+\epsilon)^2/2} |P_{>n}^{(m)}(y+\epsilon) - P_{>n}^{(m)}(y)| + |[e^{-(y+\epsilon)^2/2} - e^{-y^2/2}] P_{>n}^{(m)}(y)| \\ &\leq \epsilon e^{-(y+\epsilon)^2/2} \sup_{\alpha \in (y, y+\epsilon)} |P_{>n}^{(m+1)}(\alpha)| + \epsilon(y+\epsilon) e^{-y^2/2} |P_{>n}^{(m)}(y)| \\ &\lesssim \|X\| \epsilon [n^{(m+1)/2} + (y+\epsilon) n^{m/2}] \\ &\lesssim \epsilon n^{(m+1)/2} \|X\|. \end{aligned}$$

One could crudely estimate $P(\|X\| > e^{N^2/8}) \leq e^{-N^2/4} \mathbb{E}\|X\|^2 \lesssim n e^{-N^2/4}$. Therefore by letting $\delta = N n^{-1/2} e^{-N^2/8}$ and conditioning on the event $\|X\| \leq e^{N^2/8}$ we obtain

$$\begin{aligned} \mathbb{P}(\exists y \in I_n : |e^{-y^2/2} P_{>n}^{(m)}(y)| \gtrsim N n^{m/2}) \\ &\lesssim n N^{-1} e^{-N^2/8} + n e^{-N^2/4} \\ &\lesssim n e^{-N^2/8}. \end{aligned}$$

This completes the proof of (3.3). □

3.2. Repulsion of the real roots. In this section we prove estimates concerning the separation of real roots of P_n and P_∞ in

$$I_n = [-n^{1/2} + n^{1/6} \log n, n^{1/2} - n^{1/6} \log n].$$

LEMMA 2. For any $c_2 > 0$ the following estimates hold for $c_1 > c_2 + 2$:

- (i) $\mathbb{P}(\exists x \in I_n : P_n(x) = 0, |\frac{d}{dx}(e^{-x^2/2} P_n(x))| < n^{-c_1}) \lesssim n^{-c_2}$
- (ii) $\mathbb{P}(\exists x, x' \in I_n : P_n(x) = P_n(x') = 0, 0 < |x - x'| < n^{-c_1}) \lesssim n^{-c_2}.$

Proof. For convenience of notation let $q_n(x) = e^{-x^2/2} P_n(x)$. Clearly q_n and P_n have the same real roots. Furthermore, for $x \in I_n$ it holds that

$$\begin{aligned} q'_n(x) &= e^{-x^2/2} P'_n(x) + (-x)e^{-x^2/2} P_n(x) \\ &\leq e^{-x^2/2} |P'_n(x)| + \sqrt{n}e^{-x^2/2} |P_n(x)| \end{aligned}$$

and similarly

$$\begin{aligned} q''_n(x) &= e^{-x^2/2} P''_n(x) + 2(-x)e^{-x^2/2} P'_n(x) + (x^2 - 1)e^{-x^2/2} P_n(x) \\ &\lesssim e^{-x^2/2} |P''_n(x)| + \sqrt{n}e^{-x^2/2} |P'_n(x)| + ne^{-x^2/2} |P_n(x)|. \end{aligned}$$

Thus, using Lemma 1 with $N = C \log^{1/2} n$ with $C > 0$ large, we obtain

$$(3.5) \quad \mathbb{P}\left(\sup_{y \in I_n} |q''_n(y)| \gtrsim n \log^{1/2} n\right) \lesssim ne^{-cC \log n} < n^{-c_2}.$$

(i) Let $\delta = n^{-c_1}$.

Suppose that $q_n(x) = 0$ and $|q'_n(x)| < \delta$ for some fixed $x \in I_n$. Then for every $x' \in I_n$ with $|x' - x| \leq \delta$, conditioning on the event $\sup_{|y| \in I_n} |q''_n(y)| \lesssim n \log^{1/2} n$ and using the mean value theorem, we have

$$\begin{aligned} q_n(x') &= q_n(x) + (x' - x)q'_n(x) + O\left(\frac{(x - x')^2 n \log^{1/2} n}{2}\right) \\ &\lesssim \delta^2 + \delta^2 n \log^{1/2} n \\ &\lesssim \delta^2 n \log^{1/2} n =: \beta. \end{aligned}$$

Now, divide the interval I_n into $O(n^{1/2} \delta^{-1})$ intervals of length at most $\delta/2$. Using the above estimates and using an union bound, it follows that

$$\mathbb{P}(\exists x \in I_n : q_n(x) = 0, |q'_n(x)| < \delta) \lesssim \sqrt{n} \delta^{-1} \sup_{x \in I_n} \mathbb{P}(|q_n(x)| \lesssim \beta) + n^{-c_2}.$$

Now, for each $x \in I_n$ there is $0 \leq j \leq n$ depending on x such that $|e^{-x^2/2} \frac{x^j}{\sqrt{j!}}| \gtrsim n^{-1/2}$. To see this, we invoke (3.4) for $m = 0$ and obtain

$$e^{-x^2} \sum_{j>n} \frac{x^{2j}}{j!} = \text{Var} [e^{-x^2/2} P_{>n}(x)] \lesssim n^{1/2} e^{-\log^3 n/3}.$$

Consequently for $x \in I_n$ we have $\sum_{j=0}^n \frac{x^{2j}}{j!} \gtrsim e^{x^2}$ and therefore one could select a $j \in [0, n]$ with the stated properties.

Given such a j , we condition on $e^{-x^2/2} \sum_{i \neq j} \xi_i \frac{x^i}{\sqrt{i!}}$, which is independent from ξ_j , obtaining (for any absolute constant $C > 0$)

$$\begin{aligned} P(|q_n(x)| \leq C\beta) &\leq \sup_z P\left(\frac{x^j e^{-x^2/2}}{\sqrt{j!}} \xi_j \in (z - C\beta, z + C\beta)\right) \\ &\lesssim \frac{e^{x^2/2} \sqrt{j!}}{|x^j|} \beta \lesssim n^{1/2} \beta \end{aligned}$$

since the density of the Gaussian distribution (of ξ_j) is bounded. Note that the implicit constants are independent of $x \in I_n$. Consequently,

$$\begin{aligned} &\mathbb{P}(\exists x \in I_n : P_n(x) = 0, |P'_n(x)| < \delta) \\ &\lesssim n^{-c_2} + n\delta^{-1}\beta = n^{-c_2} + \delta n^2 \log^{1/2} n = n^{-c_2} + n^{2-c_1} \log^{1/2} n \lesssim n^{-c_2} \end{aligned}$$

provided that $c_1 > c_2 + 2$.

(ii) Assume that for some $x \neq x'$ in I_n we have $q_n(x) = q_n(x') = 0$. By the mean value theorem there is some x'' between x, x' such that $q'_n(x'') = 0$. Let $\delta = n^{-c_1}$ as before. Conditioning on the event $\sup_{|y| \in I_n} |q''_n(y)| \lesssim n \log^{1/2} n$ (which holds with probability $1 - O(n^{-c_2})$) and using the mean value theorem we have

$$q'_n(x) = q'_n(x'') + |x - x''| O(n \log^{1/2} n) = O(\delta n \log^{1/2} n)$$

therefore for any $y \in [x - \delta, x + \delta]$ it holds that

$$q_n(y) = q_n(x) + (y - x)q'_n(x) + (y - x)^2 O(n \log^{1/2} n) = O(\delta^2 n \log^{1/2} n).$$

The rest of the proof similar to (i). □

Using an entirely similar argument, we also have the following series analogue of Lemma 2.

LEMMA 3. *For any $c_2 > 0$ the following estimates hold for $c_1 > c_2 + 2$:*

- (i) $\mathbb{P}(\exists x \in I_n : P_\infty(x) = 0, |\frac{d}{dx}(e^{-x^2/2} P_\infty(x))| < n^{-c_1}) \lesssim n^{-c_2}$
- (ii) $\mathbb{P}(\exists x, x' \in I_n : P_\infty(x) = P_\infty(x') = 0, 0 < |x - x'| < n^{-c_1}) \lesssim n^{-c_2}.$

4. Local law for P_n . In this section we prove a local law for P_n , which will be used in the proof of Theorem 5 and Theorem 4.

LEMMA 4. *The following holds with overwhelming probability: for any interval $I \subset \mathbb{R}$ it holds that*

$$N_n(I) \lesssim (1 + |I \cap [-\sqrt{n}, \sqrt{n}]|) n^{o(1)}.$$

A variant of Lemma 4 for complex zeros of (non-Gaussian) Weyl polynomials was considered in [29] (see estimates (87,88) of [29]). The proof given below for Lemma 4 is inspired by the (complex) argument in [29]. Our setting is simpler because P_n is Gaussian thus our condition on I is weaker (in comparison to the requirement that $I \subset \{n^{-C} \leq |x| \leq C\sqrt{n}\}$ in [29]).

We will need the following estimate [29, Proposition 4.1, arXiv version].

PROPOSITION 1. *Let $n \geq 1$ be integer and f be a random polynomial of degree at most n . Let $z_0 \in \mathbb{C}$ be depending on n , and let $n^{-O(1)} \lesssim c \leq r \lesssim n^{O(1)}$ be quantities that may depend on n .*

Let $G : \mathbb{C} \rightarrow \mathbb{C}$ be a deterministic smooth function that may depend on n such that

$$\sup_{z \in B(z_0, r+c) \setminus B(z_0, r-c)} |G(z)| \lesssim n^{O(1)}.$$

Assume that for any $z \in B(z_0, r+c) \setminus B(z_0, r-c)$ one has

$$\log |f(z)| = G(z) + O(n^{o(1)})$$

with overwhelming probability.

Then with overwhelming probability the following holds: $f \not\equiv 0$ and the number N of roots of f in $B(z_0, r)$ satisfies

$$\begin{aligned} N &= \frac{1}{2\pi} \int_{B(z_0, r)} \Delta G(z) dz + O(n^{o(1)} c^{-1} r) \\ &\quad + O\left(\int_{B(z_0, r+c) \setminus B(z_0, r-c)} |\Delta G(z)| dz\right) \end{aligned}$$

where ΔG is the Laplacian of G .

We will use a crude estimate for the roots of P_n :

LEMMA 5. *Given any $C > 0$, with probability at least $1 - O(n^{-C})$ the roots of P_n satisfy $|z| \leq n^{(3C+2)/2}$.*

Proof. Without loss of generality assume $n \geq 2$. Let $X = (\xi_0, \dots, \xi_{n-1})$ and let $\|\cdot\|$ denote the ℓ_2 norm on $\mathbb{R}^n$. By Cauchy-Schwartz, we have the deterministic

estimate

$$\left| \sum_{j=0}^{n-1} \xi_j \frac{z^j}{\sqrt{j!}} \right| \leq \|X\| \left(\sum_{j=0}^{n-1} \frac{|z|^{2j}}{j!} \right)^{1/2}.$$

For any $|z| > n^{(3C+2)/2}$ it is clear that the sequence $(|z|^{2j}/j!)_{j=0}^n$ is lacunary

$$\frac{|z|^{2j}/j!}{|z|^{2j-2}/(j-1)!} = \frac{|z|^2}{j} \geq n^{3C+2}/j \geq n^{3C+1} > 1$$

therefore we have the deterministic bound

$$\left(\sum_{j=0}^{n-1} \frac{|z|^{2j}}{j!} \right)^{1/2} \lesssim_C n^{-(3C+1)/2} \frac{|z|^n}{\sqrt{n!}}.$$

Consequently it suffices to show that the event $\{\|X\| \leq \frac{1}{M} n^{(3C+1)/2} |\xi_n|\}$ has probability at least $1 - O_{M,C}(n^{-C})$, any $M > 0$. Since $\mathbb{E}\|X\|^2 = n$, it follows that

$$P(\|X\| < n^{(C+1)/2}) = 1 - O(n^{-C})$$

and using boundedness of the density of Gaussian we have

$$P(|\xi_n| \geq Mn^{-C}) = 1 - O_M(n^{-C})$$

thus taking the intersection of these two events we obtain the desired claim. $\square$

4.1. Proof of Lemma 4. We now begin the proof of Lemma 4. Note that $P_n/|\text{Var} P_n|^{1/2}$ is normalized Gaussian. It follows that for any z

$$\log |P_n(z)| = \frac{1}{2} \log |\text{Var} P_n(z)| + O(n^{o(1)})$$

with overwhelming probability (the implicit constant is independent of z but the bad event may depend on z). And

$$\text{Var} P_n(z) = \sum_{j=0}^n \frac{|z|^{2j}}{j!}.$$

Let z be such that $|z| \geq \sqrt{n}$. Then the sequence $1 \leq |z|^2/1! \leq \dots \leq |z|^{2n}/n!$ is increasing. It follows that $|z|^{2n}/n! \leq \text{Var}[P_n] \leq (n+1)|z|^{2n}/n!$, and consequently using Stirling's formula we have the uniform bound

$$\begin{aligned} \log |\text{Var} P_n(z)| &= 2n \log |z| - \log(n!) + O(\log n) \\ &= 2n \log |z| - (n \log n - n) + O(n^{o(1)}). \end{aligned}$$

If $|z| \leq n^{1/2}$ then $|z|^{2k}/(k)! \geq |z|^{2k+2}/(k+1)!$ for any $k \geq n$ and when $k > 2n$ we have $|z|^{2k}/k! \geq 2|z|^{2k+2}/(k+1)!$. Thus $n^{-1}e^{|z|^2} \lesssim \text{Var}[P_n] \leq e^{|z|^2}$ therefore

$$\log |\text{Var } P_n(z)| = |z|^2 + O(n^{o(1)}).$$

We now take $G(z) = \frac{1}{2}g(|z|)|z|^2 + (1 - g(|z|))[n \log |z| - \frac{1}{2}n(\log n - 1)]$ which is smooth where $g: \mathbb{R} \rightarrow [0, 1]$ a bump function such that $g(x) = 1$ for $|x| \leq \sqrt{n}$ and $g(x) = 0$ for $|x| \geq \sqrt{n} + 1$. In the transitional region $\sqrt{n} \leq |z| \leq \sqrt{n} + 1$, by examination we have

$$2n \log |z| - n \log n + n = |z|^2 + O(1).$$

Therefore for each z with overwhelming probability it holds that

$$\log |P_n(z)| = \frac{1}{2} \log |\text{Var } P_n(z)| + O(n^{o(1)}) = G(z) + O(n^{o(1)}).$$

Note that G is depending only on $|z|$ and satisfies polynomial bound $G(z) = O(n^{O(1)})$ if $|z|$ is also at most polynomial in n . Furthermore,

$$\Delta G(z) = \begin{cases} 2, & |z| \leq \sqrt{n} \\ 0, & |z| \geq \sqrt{n} + 1 \end{cases}$$

and for $\sqrt{n} < |z| < \sqrt{n} + 1$ using the polar coordinate form of Δ it holds that

$$\begin{aligned} \Delta G(z) &= \left[\frac{1}{r} \partial_r + \partial_{rr} \right] \left(g(r) \frac{r^2}{2} + (1 - g(r)) \left(n \log r - \frac{1}{2} n \log n + \frac{n}{2} \right) \right) \\ &= O\left(\frac{1}{r} |g'(r)| \right) + O(|g''(r)|) + O\left(|g'(r)| \left| r - \frac{n}{r} \right| \right) = O(1). \end{aligned}$$

Now, let $C > 0$, then by Lemma 5 with probability $1 - O(n^{-C})$ the roots of P_n satisfy $|z| \leq N := n^{(3C+2)/2}$.

We now apply Proposition 1 with $z_0 = N$, $r = N/2$, and $c = N/4$. Then with overwhelming probability

$$\begin{aligned} N_n[N/2, 3N/2] &\lesssim \int_{B(N, N/2)} 1_{B(0, \sqrt{n}+1)} + O(n^{o(1)}) \\ &\quad + \int_{B(N, 3N/2) \setminus B(N, N/2)} 1_{B(0, \sqrt{n}+1)} \\ &= O(n^{o(1)}). \end{aligned}$$

We then repeat (variance of) this argument $O(\log N)$ times with a decreasing lacunary sequence of z_0 (starting from N). Then with overwhelming probability, in $[\sqrt{n} + 2, N]$ there are $O(n^{o(1)} \log N) = O(n^{o(1)})$ real roots. By a similar argument, we have the same bound in $[-N, -\sqrt{n} - 2]$ with overwhelming probability.

We now consider the real roots in $[-\sqrt{n} - 2, \sqrt{n} + 2]$.

Let $z_0 = \sqrt{n}$ and $r = c = 2$, it follows that with overwhelming probability the number of real roots in $[\sqrt{n}-2, \sqrt{n}+2]$ is $O(n^{o(1)})$. By repeating this argument it follows that for any interval $I_0 \subset [-\sqrt{n}-2, \sqrt{n}+2]$ of length 1 with overwhelming probability the number of real roots in I is $O(n^{o(1)})$. Of course if I_0 has length less than 1 then using monotonicity of $N_n(I)$ we also have the same upper bound. Dividing $[-\sqrt{n}-2, \sqrt{n}+2]$ into intervals of length 1 and taking the union bound, it follows that one could ensure that for all subintervals of length 1 with overwhelming probability.

Consequently, given any $C > 0$, with probability $1 - O(n^{-C})$, for any interval $I \subset \mathbb{R}$ we have

$$N_n(I) \lesssim (1 + |I \cap [-\sqrt{n}, \sqrt{n}]|)n^{o(1)}.$$

This completes the proof of Lemma 4.

5. Proof of Theorem 4 assuming Theorem 5. Recall the notation that $N_n(I)$ denotes the number of real roots of P_n in $I \subset \mathbb{R}$. Let $h = 1_{[-1,1]}$ and $R_n = \sqrt{n}$. Let $N_{n,\text{in}} := N_n([-\sqrt{n}, \sqrt{n}])$ and $N_{n,\text{out}} = N_n - N_{n,\text{in}}$. Then by Theorem 5, we have

$$\begin{aligned} \text{Var}[N_{n,\text{in}}]/2\sqrt{n} &\longrightarrow K \in (0, \infty) \\ \frac{N_{n,\text{in}} - \mathbb{E}N_{n,\text{in}}}{\sqrt{\text{Var}[N_{n,\text{in}}]}} &\longrightarrow N(0, 1) \end{aligned}$$

as $n \rightarrow \infty$, and the second convergence is in distribution. By Lemma 4, with overwhelming probability we have $N_{n,\text{out}} = O(n^{o(1)})$, and we always have $N_{n,\text{out}} \leq n$ deterministically. Consequently

$$\mathbb{E}N_{n,\text{out}}^2 = O(n^{o(1)}) = o(\mathbb{E}N_{n,\text{in}}^2)$$

and therefore $\text{Var}[N_{n,\text{out}}] = O(n^{o(1)})$ and so

$$\text{Var}[N_n] = \text{Var}[N_{n,\text{in}}](1 + o(1)) = 2\sqrt{n}K(1 + o(1))$$

Furthermore, with overwhelming probability we have

$$\begin{aligned} \frac{N_n - \mathbb{E}N_n}{\sqrt{\text{Var}[N_n]}} &= o(1) + \frac{N_{n,\text{in}} - \mathbb{E}N_{n,\text{in}}}{\sqrt{\text{Var}[N_n]}} \\ &= o(1) + (N_{n,\text{in}} - \mathbb{E}N_{n,\text{in}}) \left[\frac{1}{\sqrt{\text{Var}[N_{n,\text{in}}]}} + \frac{O(n^{o(1)})}{\sqrt{\text{Var}[N_{n,\text{in}}]}\sqrt{\text{Var}[N_n]}} \right] \\ &= o(1) + \frac{N_{n,\text{in}} - \mathbb{E}N_{n,\text{in}}}{\sqrt{\text{Var}[N_{n,\text{in}}]}}(1 + o(1)). \end{aligned}$$

Thus by Slutsky's theorem (see e.g., [1, Chapter 7]) it follows that $\frac{N_n - \mathbb{E}N_n}{\sqrt{\text{Var}[N_n]}} \rightarrow N(0, 1)$ in distribution.

6. Proof of Theorem 5 assuming Theorem 6. The comparison argument in this section is inspired by similar arguments in [6, 23].

Recall that $N_n = \sum_{x \in Z_n} h(x/R_n)$ and Z_n is the multiset of the real zeros of P_n .

Denote $N_\infty := n(R, h) = \sum_{x \in Z} h(x/R_n)$ where Z is the multiset of the real zeros of P_∞ . Let $N_G = N(0, 1)$ be the standard Gaussian random variable, and

$$N_n^* := \frac{N_n - \mathbb{E}N_n}{(\text{Var } N_n)^{1/2}}, \quad N_\infty^* = \frac{N_\infty - \mathbb{E}N_\infty}{(\text{Var } N_\infty)^{1/2}}.$$

Applying Theorem 6, we obtain

$$\lim_{n \rightarrow \infty} \frac{\text{Var}[N_\infty]}{\|h\|_2^2 R_n} = K$$

and $N_\infty^* \rightarrow N_G$ in distribution.

To deduce Theorem 5, we will compare N_n with N_∞ .

LEMMA 6. *As $n \rightarrow \infty$, it holds that*

$$\mathbb{E}|N_n - N_\infty|^2 = o(R_n).$$

Below we prove Theorem 5 assuming Lemma 6.

Proof of Theorem 5. For convenience let $\Delta N_n = N_n - N_\infty$. It follows from Lemma 6 that $|\mathbb{E}\Delta N_n| = |\mathbb{E}N_n - \mathbb{E}N_\infty| = o(R_n^{1/2})$. Using the L^2 triangle inequality we also obtain $|\sqrt{\text{Var}[N_n]} - \sqrt{\text{Var}[N_\infty]}| = o(R_n^{1/2})$. Since $\text{Var}[N_\infty] = 2R_n K(1 + o(1)) = \Theta(\sqrt{n})$ as $n \rightarrow \infty$ (in other words $\text{Var}[N_\infty]$ is bounded above and below by constant multiples of $\sqrt{n}$), we obtain $\text{Var}[N_n] = 2R_n K(1 + o(1))$, in particular $\text{Var}[N_n] = \Theta(R_n)$.

Now,

$$\begin{aligned} N_n^* &= \frac{\Delta N_n - \mathbb{E}\Delta N_n}{[\text{Var } N_n]^{1/2}} + \frac{N_\infty - \mathbb{E}N_\infty}{[\text{Var } N_n]^{1/2}} \\ &= O\left(\frac{|\Delta N_n - \mathbb{E}\Delta N_n|}{R_n^{1/2}}\right) + N_\infty^* \left(1 + \frac{\sqrt{\text{Var } N_\infty} - \sqrt{\text{Var } N_n}}{\sqrt{\text{Var } N_n}}\right) \\ &= O\left(\frac{|\Delta N_n|}{R_n^{1/2}} + o(1)\right) + N_\infty^*(1 + o(1)). \end{aligned}$$

Since $\mathbb{E}|\Delta_n|^2 = o(R_n)$, it follows that $\frac{|\Delta N_n|}{R_n^{1/2}} \rightarrow 0$ in probability. Therefore by Slutsky's theorem (see e.g., [1, Chapter 7]) it follows that N_n^* converges to $N(0, 1)$ in distribution. $\square$

Our proof of Lemma 6 will use a comparison argument. More specifically, we will show that with high probability $\sup_{x \in I_n} |P_n(x) - P_\infty(x)|$ is very small in comparison to the typical distance between the real roots inside I_n of P_n and P_∞ . Via geometric considerations and properties of h , it will follow that $|N_n - N_\infty| = O(1)$ with high probability, which implies the desired estimates for $|N_n^* - N_\infty^*|$.

We will use an elementary result whose proof is elementary (see e.g., [23]).

PROPOSITION 2. *Let F and G be continuous real valued functions on $\mathbb{R}$, and $F \in C^2$. Let $\epsilon_1, M, N > 0$ and $I := [x_0 - \epsilon_1/M, x_0 + \epsilon_1/M]$. Assume that*

- $F(x_0) = 0, |F'(x_0)| \geq \epsilon_1$;
- $|F''(x)| \leq M$ for $x \in I$;
- $\sup_{x \in I} |F(x) - G(x)| \leq M'$.

Then G has a root in I if $M' \leq \frac{1}{4}\epsilon_1^2/M$.

Proof of Lemma 6. Let $h_t(x) := h(x/t)$, for $t > 0$.

Let $q_n(x) = e^{-x^2/2}P_n(x)$ and $q_\infty(x) = e^{-x^2/2}P_\infty(x)$. Note that the real roots of q_n and P_n are the same, and the real roots of q_∞ and P_∞ are the same.

Let $c_2 > 0$ and $c_1 > c_2 + 2$. Let $I_n = [-n^{1/2} + n^{1/6} \log n, n^{1/2} - n^{1/6} \log n]$ and let $J_n = \text{supp}(h_{R_n}) \setminus I_n$.

Applying Lemma 1 (with $N = C_0 \log^{1/2} n$, C_0 large), Lemma 2, Lemma 3, with probability $1 - O(n^{-c_2})$ the following event (denoted by E) holds: For every $x \in I_n$, we have

- (i) $|q_n(x) - q_\infty(x)| \leq M' := n^{-C}$
- (ii) if $q_n(x) = 0$ then $|q'_n(x)| \geq \epsilon_1 := n^{-c_1}$ and $q_n(x') \neq 0$ for all $x' \in I_n$ such that $|x - x'| \leq \epsilon_1$.
- (iii) $|q''_n(x)| \leq M := C_1 n \log^{1/2} n$, C_1 absolute constant.

By choosing $C > 2c_1 + 1$, it follows that

$$\frac{1}{4} \frac{\epsilon_1^2}{M} = \frac{n^{-2c_1-1}}{4C_1 \log n} > M' = n^{-C}.$$

Consequently, Proposition 2 applies. (Note that the zeros of P_n are at least ϵ_1 apart by (ii)). Thus for each zero of P_n in I_n (except for those near the endpoints) we could pair with one real zero of P_∞ that is within a distance $\epsilon_1/M < \epsilon_1/2$.

Similarly, we consider the event E' with $\mathbb{P}(E') \geq 1 - O(n^{-c_2})$ where the following holds: for every $x \in I_n$,

- (i) $|q_n(x) - q_\infty(x)| \leq n^{-C}$.
- (ii) If $q_\infty(x) = 0$ then $|q'_\infty(x)| \geq n^{-c_1}$ and $q'_n(x') \neq 0$ for all $x' \in I_n$ such that $|x - x'| \leq e^{-c_1 n}$.
- (iii) $|q''_\infty(x)| \lesssim n \log^{1/2} n$.

Thus by applying Proposition 2 as before it follows that on the event E' for each zero of P_∞ in I_n (except for those near the endpoints) we could pair with one real zero of P_n that is within a distance $\epsilon_1/M < \epsilon_1/2$.

Consequently, on the event $G = E' \cap E$ the zeros of P_n and P_∞ inside I_n will form pairs, except for $O(1)$ zeros near the endpoints.

Now, if $|x - x'| \leq \epsilon_1/M$ is such a pair then there are three possibilities: (i) both x and x' are inside one interval forming $\text{supp}(h)$, or (ii) both x and x' are outside $\text{supp}(h)$, or (iii) one of them is inside and one is outside.

In the last two cases we have $|h(x/R_n) - h(x'/R_n)| = O(1)$, while in the first case using Hölder continuity of h we have

$$|h(x/R_n) - h(x'/R_n)| \lesssim \left(\frac{\epsilon_1/M}{R_n} \right)^\alpha < \frac{1}{n}$$

by choosing c_2, c_1 large compare to $1/\alpha$. Since there are at most n such pairs, it follows that on the event G we have

$$|N_\infty - N_n| \lesssim 1 + M_n + M_\infty,$$

where M_n and M_∞ are the numbers of zeros of P_n and P_∞ in J_n , respectively.

Note that if $R_n \leq n^{1/2} - n^{1/6} \log n$ then $\text{supp}(h_{R_n}) \subset I_n$ and $M_n = M_\infty = 0$, so clearly $\mathbb{E}M_n^2 = \mathbb{E}M_\infty^2 = 0 = o(R_n)$. On the other hand, if $R_n > n^{1/2} - n^{1/6} \log n$ (recall that $R_n \leq n^{1/2} + o(n^{1/4})$ by given assumption) then we have $|J_n|^2 = o(n^{1/2}) = o(R_n)$. By translation invariance of the real zeros of P_∞ and using Theorem 6, it follows that

$$\mathbb{E}M_\infty^2 = O(|J_n|^2) = o(R_n).$$

By Lemma 4 we also have $M_n \leq n^{o(1)}(1 + |J_n \cap [-n^{1/2}, n^{1/2}]|) = O(n^{o(1)+1/6})$ with overwhelming probability. Since $M_n \leq n$ always, it follows that

$$\mathbb{E}M_n^2 = O(n^{1/3+o(1)}) = o(R_n).$$

Therefore, taking c_2 large we obtain

$$\begin{aligned} \mathbb{E}|N_n - N_\infty|^2 &\lesssim 1 + \mathbb{E}M_\infty^2 + \mathbb{E}M_n^2 + \mathbb{E}((N_n^2 + N_\infty^2)1_{G^c}) \\ &\lesssim o(R_n) + P(G^c)^{1/2}(\mathbb{E}[N_n^4] + \mathbb{E}[N_\infty^4])^{1/2} \\ &= o(R_n) + O(n^{-c_2/2})O(n^4) = o(R_n) \end{aligned}$$

here we have used the crude estimate $N_n = O(n)$ and the estimate $\mathbb{E}N_\infty^4 \lesssim R_n^4 = O(n^2)$ (which is a result of Theorem 6).

It follows that in both cases we have

$$\mathbb{E}|N_n - N_\infty|^2 = o(R_n).$$

This completes the proof of Lemma 6. $\square$

7. Estimates for correlation functions. In this section we summarize several new estimates for the correlation function for the real zeros of P_∞ , which will be used in the proof of Theorem 6.

We first recall the notion of correlation function. Let X be a random point process on $\mathbb{R}$. For $k \geq 1$, the function $\rho : \mathbb{R}^k \rightarrow \mathbb{R}$ is the k -point correlation function of X if for any compactly supported C^∞ function $f : \mathbb{R}^k \rightarrow \mathbb{R}$ it holds that

$$\mathbb{E} \sum_{x_1, \dots, x_k} f(x_1, \dots, x_k) = \int \dots \int_{\mathbb{R}^k} f(\xi_1, \dots, \xi_k) \rho(\xi_1, \dots, \xi_k) d\xi_1 \dots d\xi_k$$

where on the left-hand side the summation is over all ordered k -tuples of different elements in X . (In particular if $(x_\alpha)_{\alpha \in I}$ is a labeling of elements of X then we are summing over all $(x_{\alpha_1}, \dots, x_{\alpha_k})$ where $(\alpha_1, \dots, \alpha_k) \in I^k$ such that $\alpha_i \neq \alpha_j$ if $i \neq j$. The correlation function is symmetric and the definition does not depend on the choice of the labeling.) Note that this implies ρ is locally integrable on $\mathbb{R}^k$. If there is $\epsilon > 0$ such that ρ is locally $L^{1+\epsilon}$ integrable, then by a simple approximation argument it follows that the above equality holds when f is only bounded and compactly supported. In particular, for every interval $I \subset \mathbb{R}$ it holds that

$$\mathbb{E}X(I)(X(I) - 1) \dots (X(I) - k + 1) = \int \dots \int_{I^k} \rho(\xi_1, \dots, \xi_k) d\xi_1 \dots d\xi_k$$

in the above display $X(I) := |X \cap I|$.

One should point out that the k -point correlation function does not always exist (however existence of the correlation measure, generalizing $\rho(\xi_1, \dots, \xi_k) d\xi_1 \dots d\xi_k$, follows from the Riesz representation theorem). For the setting of the current work (namely P_n and P_∞), existence of the correlation functions is a consequence of the Kac-Rice formula (see also [12] for generalizations to all real Gaussian analytic functions).

Let $\rho^{(k)}$ be the k -point correlation function for the real zeros of P_∞ .

When it is clear from the context we will simply write ρ instead of $\rho^{(k)}$.

LEMMA 7. *For every $M > 0$ and $k \geq 1$ there is a finite positive constant $C_{M,k}$ such that for all $x_1, \dots, x_k \in [-M, M]$ it holds that*

$$\frac{1}{C_{M,k}} \prod_{1 \leq i < j \leq k} |x_i - x_j| \leq \rho(x_1, \dots, x_k) \leq C_{M,k} \prod_{1 \leq i < j \leq k} |x_i - x_j|.$$

Lemma 7 is a special case of the following more general result, which holds for any $2k$ -nondegenerate real Gaussian analytic functions on $\mathbb{C}$, examples include

random series $\sum_{j=0} c_j \xi_j x^j$ where ξ_j are iid normalized Gaussian, $c_0, c_1, \dots \in \mathbb{R}$ such that $\sum_j c_j^2 < \infty$ and $c_0, \dots, c_{2k-1} \neq 0$. This notion of nondegeneracy is a real variant of the complex nondegeneracy notion in [22], see Section 9 of the current paper for details.

LEMMA 8. *Let $k \geq 1$. Let f be a $2k$ -nondegenerate real Gaussian analytic function on $\mathbb{C}$. Let ρ_f denote its k -point correlation function for the real zeros. For every $M > 0$ there is a finite positive constant $C_{M,k,f}$ such that for all $x_1, \dots, x_k \in [-M, M]$ it holds that*

$$\frac{1}{C_{M,k,f}} \prod_{1 \leq i < j \leq k} |x_i - x_j| \leq \rho_f(x_1, \dots, x_k) \leq C_{M,k,f} \prod_{1 \leq i < j \leq k} |x_i - x_j|.$$

Our next estimates will be about clustering properties for ρ .

LEMMA 9. *There are finite positive constants Δ_k and C_k such that the following holds: Given any $X = (x_1, \dots, x_k)$ of distinct points in $\mathbb{R}$, for any partition $X = X_I \cup X_J$ with $d = d(X_I, X_J) \geq 2\Delta_k$ we have*

$$(7.1) \quad \left| \frac{\rho(X)}{\rho(X_I)\rho(X_J)} - 1 \right| \leq C_k \exp^{-\frac{1}{2}(d-\Delta_k)^2}.$$

Using Lemma 7, it follows that if $X = (x_1, \dots, x_k)$ splits into two clusters X_I and X_J that are sufficiently far part, then the correlation function essentially factors out. From these clustering estimates and the well-known translation invariant properties of the real zeros of P_∞ (see Lemma 20 in Appendix A for a proof), it follows that ρ is bounded globally.

LEMMA 10. *Let $\ell(t) = \min(1, |t|)$ for every $t \in \mathbb{R}$. For every $k \geq 1$ there is a finite positive constant C_k such that*

$$\frac{1}{C_k} \prod_{1 \leq i < j \leq k} \ell(|x_i - x_j|) \leq \rho(x_1, \dots, x_k) \leq C_k \prod_{1 \leq i < j \leq k} \ell(|x_i - x_j|).$$

Indeed, if $k = 1$ then the estimates hold trivially. The proof of the general case uses induction: for $k \geq 2$, if we could split $X = (x_1, \dots, x_k)$ into two groups X_1, X_2 with distance $C\Delta_k$ where C is sufficiently large (depending on k) then using Lemma 9 we have

$$\left| \frac{\rho(X)}{\rho(X_1)\rho(X_2)} - 1 \right| \leq C_k e^{-\frac{1}{2}(C-1)^2\Delta_k^2} < \frac{1}{2}$$

therefore the desired claim follows from the induction hypothesis. If no such splitting could be found then it follows from geometry that $\text{diam}(X)$ is bounded. Consequently, the desired bounds follow from the local estimates for the correlation function (Lemma 7) and the translation invariant properties of the real zeros.

Using Lemma 10 and Lemma 9, we immediately obtain the additive form of (7.1):

LEMMA 11. *There are finite positive constants Δ_k and C_k such that the following holds: Given any $X = (x_1, \dots, x_k) \in \mathbb{R}^k$, for any partition $X = X_I \cup X_J$ with $d = d(X_I, X_J) \geq 2\Delta_k$ we have*

$$(7.2) \quad |\rho(X) - \rho(X_I)\rho(X_J)| \leq C_k \exp^{-\frac{1}{2}(d - \Delta_k)^2}.$$

8. Proof of Theorem 6 using correlation function estimates. Recall that Z denotes the (multi-)set of the zeros of P_∞ and $h : \mathbb{R} \rightarrow \mathbb{R}_+$ is bounded and compactly supported, and

$$n(R, h) = \sum_{x \in Z} h(x/R)$$

for each $R > 0$. Note that $n(R, h) = n(1, h_R)$ where $h_R(x) = h(x/R)$. For convenience of notation, let $\sigma(R, h)^2$ be the variance of $n(R, h)$ and let $n^*(R; h)$ be the normalization of $n(R, h)$, namely

$$n^*(R; h) = \frac{n(R, h) - \mathbb{E}n(R, h)}{\sigma(R, h)}.$$

8.1. Bound on the moments. In this section, we will show that $\mathbb{E}[n(R, h)^k] \lesssim R^k$. Clearly it suffices to consider $h = 1_I$ for some fixed interval I . Let X denote $n(R, h)$ and let I_R denote $\{Rx : x \in I\}$, note that $|I_R| = R|I| \lesssim R$. Using the uniform bound for the correlation function of real zeros of P_∞ proved in Lemma 10, we have

$$\begin{aligned} \mathbb{E}X(X-1)\dots(X-k+1) &= \iint_{I_R \times \dots \times I_R} \rho(x_1, \dots, x_k) dx_1 \dots dx_k \\ &\lesssim |I_R|^k \min(1, \text{diam}(I_R))^{k(k-1)/2} \lesssim R^k \end{aligned}$$

then the claims follow from writing X^k as a linear combination of $X(X-1)\dots(X-j)$ with $j = 0, 1, \dots, k-1$.

8.2. Asymptotic normality. The convergence of $n^*(R, h)$ to standard Gaussian follows from the following two lemmas:

LEMMA 12. *Let $h : \mathbb{R} \rightarrow \mathbb{R}_+$ be bounded and compactly supported. Assume that there are $C, \epsilon > 0$ such that $\sigma(R, h) \geq CR^\epsilon$ for R sufficiently large. Then $n^*(R; h)$ converges in distribution to the standard Gaussian law as $R \rightarrow \infty$.*

LEMMA 13. *If $h \in L^2$ then $\sigma(R, h) \gtrsim R^{1/2}$.*

We will prove Lemma 13 in Section 8.2.1. In this section we will prove Lemma 12.

We will use the cumulant convergence theorem which will be recalled below. The cumulants $s_k(N)$ of the random variable $N = \sum_{x \in Z} h(x)$ are defined by the formal equation

$$\log \mathbb{E}(e^{\lambda N}) = \sum_{k \geq 1} \frac{s_k(N)}{k!} \lambda^k$$

in particular

$$\begin{aligned} s_1(N) &= \frac{d}{d\lambda} (\log \mathbb{E}(e^{\lambda N})) \Big|_{\lambda=0} = \mathbb{E}N \\ s_2(N) &= \frac{d^2}{d\lambda^2} (\log \mathbb{E}(e^{\lambda N})) \Big|_{\lambda=0} = \mathbb{E}N^2 - (EN)^2 = \text{Var}(N). \end{aligned}$$

The version of the cumulant convergence theorem that we use is the following result of S. Janson [14]:

THEOREM 7. (Janson) *Let $m \geq 3$. Let $X_1, X_2, \dots$ be a sequence of random variables such that as $n \rightarrow \infty$ it holds that*

- $s_1(X_n) \rightarrow 0$, and
- $s_2(X_n) \rightarrow 1$, and
- $s_j(X_n) \rightarrow 0$ for each $j \geq m$.

Then $X_n \rightarrow N(0, 1)$ in distribution as $n \rightarrow \infty$, furthermore all moments of X_n converges to the corresponding moments of $N(0, 1)$.

Since $s_1(N) \equiv \mathbb{E}N$ and $s_2(N) \equiv \text{Var}(N)$ and $n^*(R; h)$ has mean 0 and variance 1, it remains to show that the higher cumulants of $n^*(R; h)$ converge to 0 as $R \rightarrow \infty$. We will show that

LEMMA 14. *For some finite constant C_k depending only on k it holds that*

$$s_k(n(1, h)) \leq C_k \|h\|_\infty^k |\text{supp}(h)|$$

here $|\text{supp}(h)|$ is the Lebesgue measure of the support of h .

Applying Lemma 14 to $h_R(x) = h(x/R)$, it follows that $s_k(n(R, h)) \leq C_{h,k} R$. It follows from scaling symmetries and the definition of cumulants that if $N' = aN + b$ where $a > 0$ and $b \in \mathbb{R}$ are fixed constants, then $s_k(N') = a^k s_k(N)$ for any $k \geq 2$. Thus, $s_k(n^*(R, h)) = \sigma(R, h)^{-k} s_k(n(R, h))$ for all $k \geq 2$. Consequently, using the fact that $\sigma(R, h)$ grows as a positive power of R , it follows that for k sufficiently large $s_k(n^*(R; h)) \rightarrow 0$ as $R \rightarrow \infty$, as desired.

Thus it remains to prove Lemma 14. The proof uses the notion of the truncated correlation functions, whose definition is recalled below (see also Mehta [21, Section 6.1]). First, given $Z = (x_1, \dots, x_k)$ let $|Z| := k$ and let Z_I denote $(x_j)_{j \in I}$. Let

$\Pi(k)$ be the set of all partitions of $\{1, 2, \dots, k\}$ (into nonempty disjoint subsets). The truncated correlation function ρ^T is defined using the following recursive formula (see e.g., Mehta [21, Appendix A.7]):

$$(8.1) \quad \rho(Z) = \sum_{\gamma \in \Pi(k)} \rho^T(Z, \gamma)$$

here if γ is the partition $\{1, \dots, k\} = I_1 \cup \dots \cup I_j$ then $\rho^T(Z, \gamma) := \rho^T(Z_{I_1}) \dots \rho^T(Z_{I_j})$. For example, explicit computation gives $\rho^T(x_1) = \rho(x_1)$, $\rho^T(x_1, x_2) = \rho(x_1, x_2) - \rho(x_1)\rho(x_2)$.

To prove Lemma 14, we will use the following two properties:

LEMMA 15. *For any $k \geq 1$ it holds that*

$$(8.2) \quad s_k(n(1, h)) = \sum_{\gamma \in \Pi(k)} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \rho^T(x) dA(x)$$

where $|\gamma|$ is the number of subsets in the partition γ , $dA(x)$ is the Lebesgue measure on $\mathbb{R}^{|\gamma|}$, and if $\gamma_1, \dots, \gamma_j$ are the cardinality of the subsets in γ then

$$h^\gamma(x) = h(x_1)^{\gamma_1} \dots h(x_j)^{\gamma_j}.$$

LEMMA 16. *There are finite positive constants c_k, C_k such that for any $Z = (x_1, \dots, x_k)$ it holds that*

$$\rho^T(Z) \leq C_k \min(1, e^{-c_k |\text{diam}(Z)|^2}).$$

A complex variant of Lemma 15 was also considered by Nazarov-Sodin in [22], who provided a proof using a detailed algebraic computation. Lemma 15 could be proved using a similar argument, and we include a proof in Appendix B.

Proof of Lemma 16. We will use mathematical induction on k . If $k = 2$ this follows from the uniform boundedness and clustering properties of ρ :

$$|\rho^T(x_1, x_2)| = |\rho(x_1, x_2) - \rho(x_1)\rho(x_2)| \leq C \min(1, e^{-c|x_1 - x_2|^2}).$$

Let $k \geq 3$ and assume the estimates hold for all collection k' points where $1 \leq k' < k$. Then there is a partition of $Z = Z_I \cup Z_J$ based on $\{1, \dots, k\} = I \cup J$ such that $\text{dist}(Z_I, Z_J) \geq \text{diam}(Z)/Z_k$ and I and J are nonempty. It suffices to show that $|\rho^T(Z)| \leq C_k e^{-c_k d(Z_I, Z_J)^2}$.

Let $\Pi_1(k)$ be the set of non-trivial partitions of $\{1, \dots, k\}$ that mixes Z_I and Z_J , i.e., there is at least one block in the partition that intersects both Z_I and Z_J , and the partition has at least two blocks. Let $\Pi_2(k) = \Pi(k) \setminus \Pi_1(k)$. It follows

from (8.1) that

$$\rho(Z) = \rho^T(Z) + \sum_{\gamma \in \Pi_1(k)} \rho^T(Z, \gamma) + \rho(Z_I)\rho(Z_J)$$

consequently using clustering of ρ and the triangle inequality

$$|\rho^T(Z)| \leq |\rho(Z) - \rho(Z_I)\rho(Z_J)| + \sum_{\gamma \in \Pi_1(k)} |\rho^T(Z, \gamma)|.$$

Now, using the induction hypothesis and boundedness of ρ , it follows that

$$|\rho^T(Z)| \leq C_k e^{-c_k \text{dist}(Z_I, Z_J)^2} \leq C_k e^{-c_k \text{diam}(Z)^2}$$

(note that the constants c_k in different lines are not necessarily the same). □

We now finish the proof of Lemma 14. Since $|\Pi(k)| \lesssim_k 1$, using Lemma 15 it suffices to show that

$$\left| \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \rho^T(x) dA(x) \right| \leq C_k \|h\|_\infty^k |\text{supp}(h)|$$

for each $\gamma \in \Pi(k)$. Fix such a γ . Let $\gamma_1, \dots, \gamma_j$ be the length of its blocks. Using the uniform boundedness of the correlation function $\rho^{(k)}$ (Lemma 10), we obtain

$$\begin{aligned} & \left| \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \rho^T(x) dA(x) \right| \\ & \leq \int_{\mathbb{R}^j} |h(x_1)|^{\gamma_1} \dots |h(x_j)|^{\gamma_j} |\rho^T(x_1, \dots, x_j)| dx_1 \dots dx_j \\ & \leq \|h\|_\infty^k |\text{supp}(h)| \sup_{x_1 \in \mathbb{R}} \int_{\mathbb{R}^{j-1}} |\rho^T(x_1, x_2, \dots, x_j)| dx_2 \dots dx_j \\ & \leq C_k \|h\|_\infty^k |\text{supp}(h)| \end{aligned}$$

in the last estimate we used Lemma 16. □

8.2.1. Growth of the variance. In this section we prove Lemma 13. We have $\sigma(R, h) = \sigma(1, h_R)$, so we first estimate $\sigma(1, h)$ and then apply the estimate to h_R . Note that for $x \in \mathbb{R}$ we have $\rho(x) = \frac{1}{\pi}$. (See e.g., Edelman-Kostlan [8].) We then have

$$\begin{aligned} & \sigma(1, h)^2 \\ & = \iint_{\mathbb{R}^2} h(x_1)h(x_2) (\rho(x_1, x_2) - \rho(x_1)\rho(x_2) + \delta(x_1 - x_2)\rho(x_1)) dx_1 dx_2. \end{aligned}$$

Let $k(x_1, x_2) = \rho(x_1, x_2) - \rho(x_1)\rho(x_2)$, since the distribution of the real zeros is translation invariant it follows that $\rho(x_1, x_2)$ depends only on $x_1 - x_2$ (while

$\rho(x_1) = \rho(x_2) = \frac{1}{\pi}$). Thus, we may write $k(x_1, x_2) = k(x_1 - x_2)$ with

$$|k(x)| \leq C e^{-Cx^2}$$

uniformly over $x \in \mathbb{R}$, thus in particular the Fourier transform $\widehat{k} \in L^\infty \cap L^1$, and

$$\sigma(1, h)^2 = \iint_{\mathbb{R}^2} h(x_1) \overline{h(x_2)} \left[k(x_1 - x_2) + \frac{1}{\pi} \delta(x_1 - x_2) \right] dx_1 dx_2$$

where δ is the Dirac delta function. Let $*$ denote the convolution operation. Let $\widehat{h}$ be the Fourier transforms of h . Using Plancherel's identity and note that k is real valued, we obtain

$$\begin{aligned} \sigma(1, h)^2 &= \int_{\mathbb{R}} h(x_1) \overline{\left[(h * k)(x_1) + \frac{1}{\pi} (h * \delta)(x_1) \right]} dx_1 \\ &= \int_{\mathbb{R}} \widehat{h}(\xi) \overline{\left[\widehat{h}(\xi) \widehat{k}(\xi) + \frac{1}{\pi} \widehat{h}(\xi) \right]} d\xi \\ &= \int_{\mathbb{R}} |\widehat{h}(\xi)|^2 \left[\frac{1}{\pi} + \overline{\widehat{k}(\xi)} \right] d\xi = \int_{\mathbb{R}} |\widehat{h}(\xi)|^2 \left[\frac{1}{\pi} + \widehat{k}(\xi) \right] d\xi \end{aligned}$$

here in the last equality we use the fact that $\sigma(1, h)^2$ is real valued to remove the conjugate in $\widehat{k}$. Consequently,

$$\begin{aligned} \sigma(R, h)^2 &= \int_{\mathbb{R}} |R\widehat{h}(R\xi)|^2 \left[\frac{1}{\pi} + \widehat{k}(\xi) \right] d\xi \\ &= R \int_{\mathbb{R}} |\widehat{h}(u)|^2 \left[\frac{1}{\pi} + \widehat{k}\left(\frac{u}{R}\right) \right] du. \end{aligned}$$

Using $\widehat{k} \in L^\infty$ and the dominated convergence theorem, it follows that

$$\lim_{R \rightarrow \infty} \frac{\sigma(R, h)^2}{R} = \lim_{R \rightarrow \infty} \int |\widehat{h}(u)|^2 \left(\frac{1}{\pi} + \widehat{k}\left(\frac{u}{R}\right) \right) du = \left(\frac{1}{\pi} + \widehat{k}(0) \right) \|h\|_2^2.$$

Explicit computation [25] gives $\widehat{k}(0) + \frac{1}{\pi} = 0.18198\dots > 0$ (for the reader's convenience we include a self-contained derivation in Appendix C). Consequently $\sigma(R, h) \gtrsim R^{1/2}$. $\square$

9. Real Gaussian analytic functions and linear functionals. In this section we discuss real Gaussian analytic functions and linear functionals on $\mathbb{C}$. These notions are adaptations of analogous notions in [22] and will be used in Section 10 where we prove the correlation function estimates summarized in Section 7.

9.1. Real Gaussian analytic functions. We say that g is a real Gaussian analytic function (real GAF) if

$$g(z) = \sum_{j=1}^{\infty} \xi_j g_j(z)$$

where ξ_i are iid normalized real Gaussian, and $g_1, g_2, \dots$ are analytic functions on $\mathbb{C}$ such that $\sum_j |g_j|^2 < \infty$ uniformly on any compact subset of $\mathbb{C}$. In particular, uniformly over any compact subset of $\mathbb{C}$ we have $\mathbb{E}|g(z)|^2 = \sum_j |g_j(z)|^2 < \infty$.

9.2. Linear functionals. We say that L is a linear functional if for some $K \geq 1$ there are $m_1, \dots, m_K \in \mathbb{Z}$ nonnegative and $z_1, \dots, z_K \in \mathbb{C}$ and $\gamma_1, \dots, \gamma_K \in \mathbb{C}$ such that for any real GAF g it holds that

$$Lg = \sum_{j=1}^K \gamma_j g^{(m_j)}(z_j).$$

Here we require $(m_j, z_j) \neq (m_h, z_h)$ for $j \neq h$. We loosely say that z_j are the poles of L (technically speaking only the distinct elements of $\{z_j\}$ should be called the poles of L , although in this case one has to count multiplicity).

Since $\sum_j |g_j|^2 < \infty$ uniformly on compact subsets of $\mathbb{C}$, by standard arguments it follows that almost surely $\sum_{n=1}^{\infty} \xi_n g_n(z)$ converges absolutely on compact subsets of $\mathbb{C}$ to an analytic function (for a proof see e.g., [12]). Writing

$$Lg = \sum_{n \geq 1} \xi_n L(g_n)$$

and using independence of ξ_n 's, it follows that $Lg = 0$ a.s. iff $L(g_n) = 0$ for all n .

9.3. Rank of linear functions. Let L be a linear functional with poles $z_1, \dots, z_K$. Let $G \subset \mathbb{C}$ be a bounded domain with simple smooth boundary $\gamma = \partial G$ such that the poles z_j are inside the interior G° . By Cauchy's theorem, if g is analytic then

$$(9.1) \quad Lg = \int_{\gamma} g(z) r^L(z) dz, \quad r^L(z) = \frac{1}{2\pi i} \sum_{j=1}^K \frac{\gamma_j m_j!}{(z - z_j)^{m_j+1}}.$$

Now, r^L is a rational function vanishing at ∞ , and will be referred to as the kernel of L . We define the rank of L to be the degree of the denominator in any irreducible form of r^L (this notion of rank is well defined and is independent of G).

9.4. Degenerate and nondegenerate GAFs. We say that a real GAF g is d -degenerate if there is a linear functional $L \neq 0$ of rank at most d such that $Lg = 0$ almost surely. If no such linear functional exists, we say that g is d -nondegenerate.

9.5. Linear functional arises from the Kac-Rice formula. We discuss linear functionals used in the proof. Let f be a real Gaussian analytic function. Using the Kac-Rice formula for correlation functions of the real zeros of f (see e.g., [12]) asserts that: for any $(x_1, \dots, x_k) \in \mathbb{R}^k$, we have

$$(9.2) \quad \rho_k(x_1, \dots, x_k) = \frac{1}{(2\pi)^k |\det(\Gamma)|^{1/2}} \int_{\mathbb{R}^k} |\eta_1 \dots \eta_k| e^{-\frac{1}{2} \langle \Gamma^{-1} \eta, \eta \rangle} d\eta_1 \dots d\eta_k$$

where Γ is the covariance matrix of $(f(x_1), f'(x_1), \dots, f(x_k), f'(x_k))$, and

$$\eta = (0, \eta_1, \dots, 0, \eta_k) \in \mathbb{R}^{2k}.$$

Given $\gamma = (\alpha_1, \beta_1, \dots, \alpha_k, \beta_k) \in \mathbb{R}^{2k}$, via elementary computations we have

$$(9.3) \quad \langle \gamma, \Gamma \gamma \rangle = \langle \Gamma \gamma, \gamma \rangle = \mathbb{E} |Lf|^2, \quad Lf := \sum_{j=1}^k \alpha_j f(x_j) + \beta_j f'(x_j).$$

One could also define local version of L , namely for any $I \subset \mathbb{R}$ we $L_I f = \sum_{i \in I} \alpha_i f(x_i) + \beta_i f'(x_i)$ is also a linear functional. Certainly L and L_I depend on γ , however we will suppress the notational dependence for brevity, and none of the implicit constants in our estimates will depend on γ . Note that in the Kac-Rice formula, $\gamma = (0, \eta_1, \dots, 0, \eta_k)$.

9.6. Nondegeneracy of random series. Consider random infinite series $f(z) = \sum_{j \geq 0} a_j \xi_j z^j$ such that (ξ_j) are iid standard Gaussian, $a_j \in \mathbb{C}$ and $\sup_{z \in K} \sum_j |a_j z^j|^2 < \infty$ for any compact K . We now show that for such series if $a_0, \dots, a_{d-1} \neq 0$ then f is d -nondegenerate on $\mathbb{C}$. (Certainly f is a real GAF.)

Assume towards a contradiction that f is d -degenerate. Then there is a linear functional L of rank at most d such that $L(a_n z^n) = 0$ for all $n \geq 0$. Since $a_n \neq 0$ for $0 \leq n \leq d-1$, it follows that $L(z^n) = 0$ for all $0 \leq n \leq d-1$. Taking $\gamma = \{|z| = R\}$ for any $R > 0$ sufficiently large so that the poles of L are enclosed inside γ , we get

$$0 = L(z^n) = \int_{\gamma} z^n r^L(z) dz$$

for all $n \in \overline{0, d-1}$. Since rank of L is at most d there is some $m \in \{1, d\}$ and $C \neq 0$ such that $z^m r^L(z) = C(1 + o(1))$ as $|z| \rightarrow \infty$ uniformly. Consequently,

$$\int_{|z|=R} z^{m-1} r^L(z) dz = \int_{\gamma} z^m r^L(z) \frac{dz}{z} \longrightarrow 2\pi i C \neq 0$$

as $R \rightarrow \infty$ contradiction.

It follows from the above discussion that the infinite flat series P_{∞} is $2k$ -nondegenerate, and the Gaussian Kac polynomial of degree $2k-1$ defined by $g_{2k-1}(x) = \xi_0 + \xi_1 x + \dots + \xi_{2k-1} x^{2k-1}$ is also $2k$ -nondegenerate.

9.7. Equivalence of linear functionals. The following lemma is a real Gaussian adaptation of a result in [22].

LEMMA 17. Assume that f is d -nondegenerate real GAF. Let $K \subset \mathbb{C}$ be nonempty compact. Let G be a bounded domain such that $K \subset G^\circ$, and assume that $\gamma = \partial G$ is a simple rectifiable curve.

Then for any $d \geq 1$ there is a finite positive constant $C = C(d, G, K, f)$ such that for every linear functional L of rank at most d with poles in K we have

$$\frac{1}{C} \max_{z \in \gamma} |r^L(z)|^2 \leq \mathbb{E}|Lf|^2 \leq C \max_{z \in \gamma} |r^L(z)|^2.$$

Proof. We largely follow [22].

We first show the upper bound. Let ds denote the arc length measure on γ , then using (9.1) and Cauchy-Schwarz we have

$$\begin{aligned} |Lf|^2 &\leq \text{length}(\gamma) \max_{z \in \gamma} |r^L(z)|^2 \int_{\gamma} |f(z)|^2 ds, \\ \mathbb{E}|Lf|^2 &\leq \left(\text{length}(\gamma) \max_{z \in \gamma} |r^L(z)|^2 \right) \int_{\gamma} \mathbb{E}|f|^2 \leq C_{\gamma, f} \max_{\gamma} |r^L|^2. \end{aligned}$$

We now show the lower bound. Assume towards a contradiction that the lower bound does not hold, then there is a sequence $(L_n)_{n \geq 1}$ of linear functionals of rank at most d (with poles in K) such that $\max_{\gamma} |r_n| = 1$ but

$$\lim_{n \rightarrow \infty} \mathbb{E}|L_n f|^2 = 0.$$

We write $r_n(z) = \frac{p_n(z)}{q_n(z)}$ where p_n and q_n are polynomials, and by multiplying both the numerator and denominator of r_n by common factors (of the form $(x - \alpha)$ with $\alpha \in K$) if necessary we may assume that $\deg(q_n) = d$ and $\deg(p_n) \leq d - 1$ and q_n is monic. Since the zeros of q_n are in K , we have $\sup_{z \in \gamma} |q_n(z)| < C_{d, K} < \infty$ (uniformly over n), therefore using $\sup_{\gamma} |r_n(z)| \leq 1$ we obtain $\sup_{z \in \gamma} |p_n(z)| < C_{d, K}$ uniformly over n . Therefore, by passing to a subsequence, we may assume that (p_n) converges uniformly on γ to p . Now, using Szegő's theorem [28] (see also the survey [24]) we have

$$\sup_{z \in \gamma} |p'_n(z)| \lesssim_{\deg(p_n), \gamma} C_{d, K} = O(1).$$

Thus we could again pass to a subsequence and obtain uniform convergence for (p'_n) on γ , which will converge to p' . By iteratively passing to subsequences we may assume further that $p'_n, p''_n, \dots, p_n^{(d)}$ converge uniformly on γ to $p', p'', \dots, p^{(d)}$. Since $\deg(p_n) < d$, it follows that $p^{(d)} \equiv 0$ and consequently p is a polynomial of degree at most $d - 1$.

Now, the d complex zeros of q_n are in K , a compact set, therefore by passing to a subsequence we may assume that uniformly on $\gamma = \partial G$ we have $q_n \rightarrow q$, and q is a monic polynomial of degree d with zeros in K .

Using partial fractional decomposition of $r(z) = p(z)/q(z)$, we obtain a linear functional L of rank at most d with poles in K such that $\max_{z \in \gamma} |r^L(z) - r^{L_n}(z)| \rightarrow 0$ when $n \rightarrow \infty$. Consequently using the upper bound (already shown above) we obtain

$$\mathbb{E}|L_n f - Lf|^2 = O\left(\max_{\gamma} |r^{L_n} - r^L|^2\right) = o(1).$$

Using $\lim_{n \rightarrow \infty} \mathbb{E}|L_n f|^2 = o(1)$ it follows that $\mathbb{E}|Lf|^2 = 0$, hence $Lf = 0$ almost surely. This violates the d -nondegeneracy of f . $\square$

10. Proof of the correlation function estimates in Section 7.

10.1. Local estimates for correlation functions. In this section we prove Lemma 7 and Lemma 8. Using Lemma 17 and the Kac-Rice formula (9.2), we observe that if f_1 and f_2 are two $2k$ -nondegenerate real Gaussian analytic functions and $\rho^{[1]}$ and $\rho^{[2]}$ are the corresponding k -point correlation functions for the real zeros, then there is a finite positive constant $C = C_{M,N,k,f_1,f_2}$ such that

$$\frac{1}{C} \rho^{[2]}(y_1, \dots, y_k) \leq \rho^{[1]}(y_1, \dots, y_k) \leq C \rho^{[2]}(y_1, \dots, y_k).$$

Indeed, let Γ_j be the covariance matrix for $(f_j(y_1), f'_j(y_1), \dots, f_j(y_k), f'_j(y_k))$, which are positive definite symmetric. Then by Lemma 17, it follows that $\det(\Gamma_1)$ and $\det(\Gamma_2)$ are comparable and $\langle \Gamma_1^{-1} \eta, \eta \rangle$ and $\langle \Gamma_2^{-1} \eta, \eta \rangle$ are comparable. Consequently, using the Kac-Rice formula (9.2) it follows that $\rho^{[1]}(y_1, \dots, y_k)$ and $\rho^{[2]}(y_1, \dots, y_k)$ are comparable.

Therefore it suffices to show Theorem 7. Namely, we will show that the correlation function for the real zeros of P_∞ is locally comparable to the Vandermonde product.

Let $M > 0$ and $k \geq 1$. Assume that $x_1, \dots, x_k \in [-M, M]$. Let $N = N(M, k)$ be a large positive constant that will be chosen later. Thanks to the translation invariant property of the distribution of real zeros of Z , we have $\rho(x_1, \dots, x_k) = \rho(x_1 + N, \dots, x_k + N)$. Let $y_1 = x_1 + N, \dots, y_k = x_k + N$. Then $N - M \leq y_j \leq N + M$, and our choice of N will ensure in particular that $N - M$ and $N + M$ are very large.

We now apply the above observation to $f_1 = P_\infty$ and $f_2 = g_{2k-1} := \xi_0 + \xi_1 x + \dots + \xi_{2k-1} x^{2k-1}$, the Gaussian Kac polynomial. It then suffices to show that for any $n \geq k$ the correlation function ρ_{Kac} for the real zeros of the Gaussian Kac

polynomial $g_n(x) = \xi_0 + \xi_1 x + \cdots + \xi_n x^n$ satisfies

$$\rho_{\text{Kac}}(y_1, \dots, y_k) \sim_{M,N,k,n} \prod_{1 \leq i < j \leq k} |y_i - y_j|$$

whenever $y_1, \dots, y_k \in [N - M, N + M]$ and $N - M \gg 1$.

We now observe that the distribution of the real roots of the Kac polynomial g_n is invariant under the transformation $x \mapsto 1/x$. Indeed, $g_n(1/x) = x^{-n} \tilde{g}_n(x)$ and $\tilde{g}_n = \xi_n + \xi_{n-1}x + \cdots + \xi_0 x^n$ has the same distribution as $g_n(x)$.

It follows that, with $w_j = 1/y_j$,

$$\rho_{\text{Kac}}(y_1, \dots, y_k) \sim_{M,N,k,n} \rho_{\text{Kac}}(w_1, \dots, w_k).$$

Indeed, using the Lebesgue differentiation theorem, we have

$$\rho_{\text{Kac}}(y_1, \dots, y_k) = \lim_{\epsilon \rightarrow 0} \frac{1}{(2\epsilon)^k} \int_{B_\epsilon(y_1) \times \cdots \times B_\epsilon(y_k)} \rho_{\text{Kac}}(x_1, \dots, x_k) dx_1 \dots dx_k$$

and using the definition of correlation function this is the same as

$$= \lim_{\epsilon \rightarrow 0} \frac{\mathbb{P}(|u_1 - y_1| \leq \epsilon, \dots, |u_k - y_k| \leq \epsilon \mid g_n(u_1) = \cdots = g_n(u_k) = 0)}{(2\epsilon)^k}$$

here in the limit we condition on the event that $(u_1, \dots, u_k)$ is a k -tuple of real zeros of g_n .

Now, observe that if $|u_1 - y_1| \leq \epsilon$ and $\epsilon > 0$ is sufficiently small then $|\frac{1}{u_1} - \frac{1}{y_1}| \leq \epsilon$ where the implicit constant depends on M, N . Conversely if $|\frac{1}{u_1} - \frac{1}{y_1}| \leq \epsilon/C$ for C very large depending on M, N then for $\epsilon > 0$ sufficiently small we will have $|u_1 - y_1| \leq \epsilon$. It follows that $\rho_{\text{Kac}}(y_1, \dots, y_k)$ is comparable to the limit

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0} \frac{\mathbb{P}(|\frac{1}{u_1} - w_1| \leq \epsilon, \dots, |\frac{1}{u_k} - w_k| \leq \epsilon \mid g_n(u_1) = \cdots = g_n(u_k) = 0)}{(2\epsilon)^k} \\ &= \lim_{\epsilon \rightarrow 0} \frac{\mathbb{P}(|\frac{1}{u_1} - w_1| \leq \epsilon, \dots, |\frac{1}{u_k} - w_k| \leq \epsilon \mid \tilde{g}_n(\frac{1}{u_1}) = \cdots = \tilde{g}_n(\frac{1}{u_k}) = 0)}{(2\epsilon)^k} \end{aligned}$$

which is exactly $\rho_{\text{Kac}}(w_1, \dots, w_k)$ using the Lebesgue differentiation theorem again and the fact that $\tilde{g}_n$ has the same distribution as g_n .

Now, note that we also have

$$\prod_{1 \leq i < j \leq k} |w_i - w_j| \sim_{M,N,k,n} \prod_{1 \leq i < j \leq k} |y_i - y_j|$$

and note that $|w_j| \leq \frac{1}{N-M}$ which could be made small if N is chosen large. Therefore it suffices to show that for $\delta > 0$ sufficiently small depending on k and n there

is a finite positive constant $C = C_{\delta,k,n}$ such that for any $w_1, \dots, w_k \in [-\delta, \delta]$ it holds that

$$\frac{1}{C} \leq \frac{\rho_{\text{Kac}}(w_1, \dots, w_k)}{\prod_{1 \leq i < j \leq k} |w_i - w_j|} \leq C.$$

To show this estimate, our starting point is an explicit formula due to Gotze-Kaliada-Zaporozhets [10, Theorem 2.3] for the real correlation of the general random polynomial

$$f(x) = \gamma_0 + \gamma_1 x + \dots + \gamma_n x^n$$

where γ_j 's are independent real-valued random variables, and the distribution of γ_j has probability density f_j . To formulate the formula, we first fix some notations. Given $w = (w_1, \dots, w_k)$ and $0 \leq i \leq k$ we define $\sigma_i(w)$ to be the i th symmetric function of x , namely the sum of all products of i coordinates of w :

$$\sigma_i(w) = \sum_{1 \leq j_1 < j_2 < \dots < j_i \leq k} w_{j_1} \dots w_{j_i}$$

(if $i > k$ or $k < 0$ then $\sigma_i := 0$). Then we have, using [10, Theorem 2.3],

$$\begin{aligned} \rho(w_1, \dots, w_k) &= \prod_{1 \leq i < j \leq k} |w_i - w_j| \int_{\mathbb{R}^{n-k+1}} \prod_{i=0}^n f_i \left(\sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(w) t_j \right) \\ &\quad \times \prod_{i=1}^k \left| \sum_{j=0}^{n-k} t_j w_i^j \right| dt_0 \dots dt_{n-k}. \end{aligned}$$

We apply this to $f = g_n$ the Gaussian Kac polynomial of degree n , note that $f_j(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \lesssim 1$.

Note that if $\max |w_i| \leq \delta$ for δ very small then for $k \leq i \leq n$ we have

$$\sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(w) t_j = t_{i-k} + O(\delta \max_j |t_j|)$$

therefore with δ small enough (depending on k and n)

$$\sum_{i=k}^n \left| \sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(x) t_j \right|^2 \sim_{n,k} \sum_{j=0}^{n-k} t_j^2.$$

Since $0 \leq |\sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(x) t_j|^2 \lesssim_{k,n} \sum_{j=0}^{n-k} t_j^2$ for any i (in particular for those $0 \leq i < k$), it follows immediately that

$$\sum_{i=0}^n \left| \sum_{j=0}^k (-1)^{k-i+j} \sigma_{k-i+j}(x) t_j \right|^2 \sim_{n,k} \sum_{j=0}^{n-k} t_j^2.$$

Therefore for some finite positive constants C_1, C_2 that may depend on k, n it holds that

$$e^{-C_1 \sum_{j=0}^{n-k} t_j^2} \lesssim_{k,n} \prod_{i=0}^n f_i \left(\sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(x) t_j \right) \lesssim_{k,n} e^{-C_2 \sum_{j=0}^{n-k} t_j^2}.$$

From here it follows easily that

$$\rho(x_1, \dots, x_k) \lesssim_{\delta, k, n} \prod_{1 \leq i < j \leq k} |x_i - x_j|.$$

For the lower bound, note that if $t_1, \dots, t_{n-k} \in [-1, 1]$ then

$$\prod_{i=1}^k \left| \sum_{j=0}^{n-k} t_j w_i^j \right| = |t_0|^k + O_{k,n}(\delta),$$

so

$$\begin{aligned} & \int_{\mathbb{R}^{n-k+1}} \prod_{i=0}^n f_i \left(\sum_{j=0}^{n-k} (-1)^{k-i+j} \sigma_{k-i+j}(w) t_j \right) \prod_{i=1}^k \left| \sum_{j=0}^{n-k} t_j w_i^j \right| dt_0 \dots dt_{n-k} \\ & \gtrsim_{k,n} \int_{\mathbb{R} \times [-1, 1]^{n-k}} (|t_0|^k + O(\delta)) e^{-C_1(t_0^2 + \dots + t_{n-k}^2)} dt_0 \dots dt_{n-k} \\ & \gtrsim \int_{\mathbb{R}} |t_0|^k e^{-C_1 t_0^2} dt_0 + O(\delta) \gtrsim_{k,n} 1 \end{aligned}$$

if δ is sufficiently small.

This completes the proof of Lemma 7. $\square$

10.2. Clustering estimates for correlation functions. For a set $X = \{x_1, \dots, x_k\}$ of k distinct points and for any nonempty subset $I \subset \{1, \dots, k\}$, we denote by X_I the corresponding subset $\{x_i : i \in I\}$. Recall that ρ denote the correlation function of the real zeros of P_∞ . For simplicity of notation in this section let $f = P_\infty$.

In this section we will prove Lemma 9, namely we will show that there is a constant $\Delta_k > 0$ and C_k finite such that the following holds: Given any $X = (x_1, \dots, x_k)$ of distinct points in $\mathbb{R}$, for any partition $X = X_I \cup X_J$ with $d =$

$d(X_I, X_J) \geq 2\Delta_k$ we have

$$\left| \frac{\rho(X)}{\rho(X_I)\rho(X_J)} - 1 \right| \leq C_k \exp^{-\frac{1}{2}(d-\Delta_k)^2}.$$

We will need the following lemma. Below fix $\eta = (0, \eta_1, \dots, 0, \eta_k)$ where $\eta_1, \dots, \eta_k \in \mathbb{R}$, none of the implicit constants will depend on η . Let the linear functionals L be defined using

$$(10.1) \quad Lf = \sum_{1 \leq j \leq k} \eta_j f'(x_j)$$

and define $L_I f$ for any $I \subset \{1, \dots, k\}$ using the summation over $j \in I$ instead of $1 \leq j \leq k$.

LEMMA 18. *There are finite positive constants Δ_k and C_k such that the following holds: Given any $X = (x_1, \dots, x_k)$ of distinct points in $\mathbb{R}$, for any partition $X = X_I \cup X_J$ with $d = d(X_I, X_J) \geq 2\Delta_k$ we have*

$$|\mathbb{E}(L_I f)(L_J f)| \leq C_k e^{-\frac{1}{2}(d-\Delta_k)^2} \left(\mathbb{E}|L_I f|^2 + \mathbb{E}|L_J f|^2 \right).$$

We defer the proof of this lemma to later sections. Below we prove the clustering property of the correlation function using this lemma.

Let C_k and Δ_k be as in Lemma 18. Let $\epsilon = \frac{1}{2}C_k e^{-\frac{1}{2}(d-\Delta_k)^2}$ where $d = d(X_I, X_J)$. To show clustering it suffices to show that

$$(10.2) \quad \left(\frac{1-\epsilon}{1+\epsilon} \right)^k \rho(X_I)\rho(X_J) \leq \rho(X) \leq \left(\frac{1+\epsilon}{1-\epsilon} \right)^k \rho(X_I)\rho(X_J).$$

Define L using (10.1). By Lemma 18 we have

$$|\mathbb{E}(L_I f)(L_J f)| \leq \frac{1}{2}\epsilon \left(\mathbb{E}|L_I f|^2 + \mathbb{E}|L_J f|^2 \right)$$

therefore

$$(1-\epsilon)(\mathbb{E}|L_I f|^2 + \mathbb{E}|L_J f|^2) \leq \mathbb{E}|L f|^2 \leq (1+\epsilon)(\mathbb{E}|L_I f|^2 + \mathbb{E}|L_J f|^2).$$

Consequently using (9.3) we obtain

$$(10.3) \quad (1-\epsilon)\Gamma_{I,J} \leq \Gamma \leq (1+\epsilon)\Gamma_{I,J}$$

where Γ is the covariance matrix of $(f(x_1), f'(x_1), \dots, f(x_k), f'(x_k))$ and $\Gamma_{I,J} = \begin{pmatrix} \Gamma_I & 0 \\ 0 & \Gamma_J \end{pmatrix}$.

We obtain $\det(\Gamma) \geq (1 - \epsilon)^{2k} \det(\Gamma_I) \det(\Gamma_J)$, and therefore

$$\begin{aligned} \rho(X) &\leq (1 - \epsilon)^{-k} \frac{1}{(2\pi)^k |\det(\Gamma_{I,J})|^{1/2}} \int_{\mathbb{R}^k} |\eta_1 \dots \eta_k| e^{-\frac{1}{2}(1+\epsilon)^{-1} \langle \Gamma_{I,J}^{-1} \eta, \eta \rangle} d\eta \\ &= (1 - \epsilon)^{-k} (1 + \epsilon)^k \frac{1}{(2\pi)^{|I|} |\det(\Gamma_I)|^{1/2}} \int_{\mathbb{R}^{|I|}} |\eta_1 \dots \eta_{|I|}| e^{-\frac{1}{2}(1+\epsilon)^{-1} \langle \Gamma_I^{-1} \eta, \eta \rangle} d\eta \\ &\quad \times \frac{1}{(2\pi)^{|J|} |\det(\Gamma_J)|^{1/2}} \int_{\mathbb{R}^{|J|}} |\eta_1 \dots \eta_{|J|}| e^{-\frac{1}{2}(1+\epsilon)^{-1} \langle \Gamma_J^{-1} \eta, \eta \rangle} d\eta \\ &= \left(\frac{1 + \epsilon}{1 - \epsilon} \right)^k \rho(X_I) \rho(X_J). \end{aligned}$$

Similarly we have

$$\rho(X) \geq \left(\frac{1 - \epsilon}{1 + \epsilon} \right)^k \rho(X_I) \rho(X_J).$$

This completes the proof of (10.2). □

10.2.1. Proof of Lemma 18. We first show a small scale version of the lemma.

LEMMA 19. *Let $\rho > 0$. Suppose that K_1 and K_2 are two intervals with length at most 2ρ . Assume that L_{K_j} is a linear functional on $\mathbb{C}$ with poles inside K_j with rank at most k . Assume that $d = \text{dist}(K_1, K_2) \geq 2\rho$. Then*

$$\mathbb{E}(L_{K_1} f)(L_{K_2} f) \leq C_{f,k,\rho} e^{-\frac{1}{2}(d-2\rho)^2} (\mathbb{E}|L_{K_1} f|^2 + \mathbb{E}|L_{K_2} f|^2).$$

Proof. Let c_1 and c_2 be the centers of K_1 and K_2 . Let $T_x f$ denote

$$T_x f(z) = f(z + x) e^{-(\frac{1}{2}|x|^2 + xz)}$$

then $\mathbb{E} T_x f(t_1) T_x f(t_2) = \mathbb{E} f(t_1) f(t_2)$ for any $t_1, t_2 \in \mathbb{R}$. Therefore for any $x \in \mathbb{R}$, $T_x f$ and f have the same distribution (in particular the distribution of the real zeros of f is translation invariant). Let K_3 and K_4 be $K_1 - c_1$ and $K_4 = K_2 - c_2$, thus these intervals are centered at 0 and have length at most 2ρ . Let L_1 be such that $L_1 T_{c_1} f = L_{K_1} f$ and let L_2 be such that $L_2 T_{c_2} f = L_{K_2} f$. Then it is clear that L_1 and L_2 are linear functional of rank at most k with poles inside $[-\rho, \rho]$. Let $\gamma = \{|z| = 2\rho\}$. Since f is $2k$ -nondegenerate on $\mathbb{C}$, we then have

$$\begin{aligned} \mathbb{E}(L_{K_1} f L_{K_2} f) &= \mathbb{E}(L_1 T_{c_1} f)(L_2 T_{c_2} f) \\ &\leq \left| \int_{\gamma} \int_{\gamma} r^{L_1}(z_1) r^{L_2}(z_2) \mathbb{E}(T_{c_1} f)(z_1) T_{c_2} f(z_2) dz_1 dz_2 \right| \\ &\leq C_{\rho} \sup_{z_1, z_2 \in \gamma} |\mathbb{E}(T_{c_1} f(z_1) T_{c_2} f(z_2))| \left(\max_{\gamma} |r^{L_1}|^2 + \max_{\gamma} |r^{L_2}|^2 \right). \end{aligned}$$

By explicit computation and using $|c_1 - c_2| \geq 2\rho$, we obtain

$$\sup_{z_1, z_2 \in \gamma} \mathbb{E}(T_{c_1} f(z_1) T_{c_2} f(z_2)) \leq C_\rho e^{-(|c_1 - c_2| - 2\rho)^2}$$

therefore

$$\begin{aligned} \mathbb{E}(L_{K_1} f L_{K_2} f) &\leq C_\rho e^{-(|c_1 - c_2| - 2\rho)^2} \left(\max_{\gamma} |r^{L_1}|^2 + \max_{\gamma} |r^{L_2}|^2 \right) \\ &\leq C_{f,k,\rho} e^{-|d-2\rho|^2} \left(\mathbb{E}|L_1 f|^2 + \mathbb{E}|L_2 f|^2 \right). \end{aligned}$$

Since f and $T_{c_j} f$ have the same distribution, the last display is the same as

$$= C_{f,k,\rho} e^{-|d-2\rho|^2} \left(\mathbb{E}|L_{K_1} f|^2 + \mathbb{E}|L_{K_2} f|^2 \right)$$

which implies the desired estimate. $\square$

We now start the proof of Lemma 18. We will construct a covering X by small intervals having the following properties:

- (i) The cover will consists of $m \leq k$ intervals $I_1, \dots, I_m$ each of length at most ρ such that the distance between the centers of any two of them is at least 4ρ .
- (ii) The algorithm will ensure that $\rho > 1$ (or any given large absolute constant) but $\rho = O_{k,f}(1)$.

We first let $\rho_1 = 1$ and use the given points as centers of the interval.

If there are two centers with distance not larger than $4\rho_1$, we replace these two centers by one center at their midpoint, and enlarge all intervals, replacing ρ_1 by $\rho_2 = 3\rho_1$.

We repeat this process if needed, and since there are only k points the process has to stop. Clearly the last radius is at most 3^{k-1} .

Note that we could ensure that ρ is larger than any given constant depending on k, f if needed, by setting ρ_1 the initial radius to be larger than this constant.

Now, choose $\Delta_k > 4\rho$ such that $\Delta_k - 4\rho$ is very large compared to 1.

Notice that there is no k such that I_k intersects both I and J . Let $A = \{k : I_k \cap I \neq \emptyset\}$ and $B = \{k : I_k \cap J \neq \emptyset\}$. Now using the above small scale result we have

$$\mathbb{E}|L_I f|^2 = \sum_{k \in A} \mathbb{E}|L_{I_k} f|^2 + \sum_{k, n \in A: k \neq n} \mathbb{E}(L_{I_k} f L_{I_n} f) \geq \frac{1}{2} \sum_{k \in A} \mathbb{E}|L_{I_k} f|^2.$$

Similarly, $\mathbb{E}|L_J f|^2 \geq \frac{1}{2} \sum_{k \in B} \mathbb{E}|L_{I_k} f|^2$.

Now, if $k \in A$ and $n \in B$ it is clear that $\text{dist}(I_k, I_n) \geq d - 2\rho > 2\rho$ since $d \geq 2\Delta_k$ is very large compared to ρ . Therefore

$$\begin{aligned} |\mathbb{E}(L_I f)(L_J f)| &\leq C_{f,k,\rho} \sum_{k \in A, n \in B} e^{-\frac{1}{2}|d-4\rho|^2} \left(\mathbb{E}|L_{I_k} f|^2 + \mathbb{E}|L_{I_n} f|^2 \right) \\ &\leq C_{f,k,\rho} e^{-\frac{1}{2}|d-4\rho|^2} \left(\mathbb{E}|L_I f|^2 + \mathbb{E}|L_J f|^2 \right). \quad \square \end{aligned}$$

Appendix A. Translation invariance of the real zeros of P_∞ . The following property is standard, we include a proof for the convenience of the reader.

LEMMA 20. *The distribution of the real roots of P_∞ is invariant under translations on $\mathbb{R}$ and the reflection $x \mapsto -x$.*

Proof. Notice that $f(x)$ is a real Gaussian process with correlation function

$$K_f(x, y) = \mathbb{E}f(x)f(y) = \sum_{k \geq 0} \frac{x^k y^k}{k!} = e^{xy}$$

Let $g(x) = e^{-bx + \frac{1}{2}b^2} f(ax + b)$ where $a \in \{-1, 1\}$ and $b \in \mathbb{R}$. Then g is also a centered real Gaussian process with the correlation function

$$\begin{aligned} K_g(x, y) &= K_f(ax + b, ay + b) = e^{(ax+b)(ay+b) - b(ax+b) - b(ay+b) + b^2} \\ &= e^{xy} = K_f(x, y). \end{aligned}$$

It follows that g has the same distribution as $f(x)$. Consequently the real zeros of $f(ax + b)$ has the same distribution as the real zeros of $f(x)$. $\square$

Appendix B. Relation between cumulants and truncated correlation functions. For the convenience of the reader, we include a self-contained proof of Lemma 15 in this section, which largely follows an argument in Nazarov-Sodin [22]. Recall that X is the random point process for the real zeros of P_∞ , $h : \mathbb{R} \rightarrow \mathbb{R}_+$ is bounded compactly supported, and $n(1, h) = \sum_{\alpha \in X} h(\alpha)$, and $\Pi(k)$ is the collection of all partition of $\{1, \dots, k\}$ into disjoint nonempty subsets, and for each $\gamma \in \Pi(k)$ let $|\gamma|$ be the number of subsets in the partition γ , and if $\gamma_1, \dots, \gamma_j$ are the cardinality of the subsets in γ then

$$h^\gamma(x) = h(x_1)^{\gamma_1} \dots h(x_j)^{\gamma_j}.$$

We first prove an analogous relation between the moment and the (standard) correlation functions: if $m_k(N) = \mathbb{E}N^k$ denotes the k th moment of the random variable N , then

$$(B.1) \quad m_k(n(1, h)) = \sum_{\gamma \in \Pi(k)} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \rho(x) dx.$$

Indeed, divide $\Pi(k)$ into $\bigcup_{j \geq 1} \Pi(k, j)$ basing on $|\gamma| = j$. For each $\gamma \in \Pi(k, j)$ let $A_1, \dots, A_j$ be the subsets in the partition and let $\gamma_j = |A_j|$. Note that the correlation functions are uniformly bounded therefore we may use the bounded compactly supported function $H(x_1, \dots, x_j) = h(x_1)^{\gamma_1} \dots h(x_j)^{\gamma_j}$ as a test function in the defining property of correlation functions, and obtain

$$\begin{aligned} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) dx &= \int_{\mathbb{R}^j} h(x_1)^{\gamma_1} \dots h(x_j)^{\gamma_j} \rho(x_1, \dots, x_j) dx_1 \dots dx_j \\ &= \mathbb{E} \sum_{(\xi_1, \dots, \xi_j)} h(\xi_1)^{\gamma_1} \dots h(\xi_j)^{\gamma_j} \end{aligned}$$

where the last summation is over all ordered j -tuples of different elements of X . By summing over all possible values of j and $\gamma_1, \dots, \gamma_j \geq 1$ (with $\gamma_1 + \dots + \gamma_j = k$), it follows that

$$\begin{aligned} \sum_j \sum_{\gamma \in \Pi(k, j)} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) dx &= \mathbb{E} \sum_j \sum_{\gamma_1 + \dots + \gamma_j = k} \sum_{(\xi_1, \dots, \xi_j)} h(\xi_1)^{\gamma_1} \dots h(\xi_j)^{\gamma_j} \\ &= \mathbb{E} \left[\sum_{\xi \in X} h(\xi) \right]^k = m_k(n(1, h)). \end{aligned}$$

We now use induction on k to prove (8.2). Clearly (8.2) holds for $k = 1$. The key ingredient for the induction step is the following relationship between cumulant and moments (see e.g., [26, Chapter 2])

$$s_k = m_k - \sum_{j \geq 2} \sum_{\pi \in \Pi(k, j)} s_{\pi_1} \dots s_{\pi_j}$$

which is analogous to the following reformulation of (8.1):

$$\rho^T(Z) = \rho(Z) - \sum_{j \geq 2} \sum_{\gamma \in \Pi(k, j)} \rho^T(Z, \gamma)$$

where $\rho^T(Z, \gamma) = \rho^T(Z_{\Gamma_1}) \dots \rho^T(Z_{\Gamma_j})$ if $\gamma = (\Gamma_1, \dots, \Gamma_j)$.

To facilitate the notation, for $\gamma, \pi \in \Pi(k)$ we say that $\gamma \leq \pi$ if γ is a refinement of π , in other words the partitioning subsets of γ are subsets of the partitioning subsets in π . If $\gamma \leq \pi$ and $\gamma \neq \pi$ we say $\gamma < \pi$.

Let 1 denote the trivial partition with just one partitioning subset, clearly all $\pi \in \Pi(k)$ satisfies $\pi \ll 1$. We will write $\pi = (\Pi_1, \dots, \Pi_{|\pi|})$ below.

Using the induction hypothesis we have

$$\begin{aligned} s_k &= m_k - \sum_{\pi < 1} \prod_{j=1}^{|\pi|} s_{\pi_j} \\ &= \sum_{\gamma \leq 1} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \rho(x) dx \\ &\quad - \sum_{\pi < 1} \prod_{j=1}^{|\pi|} \sum_{\Gamma(j) \text{ partitions } \Pi_j} \int_{\mathbb{R}^{|\Gamma(j)|}} h^{\Gamma(j)}(x^{(j)}) \rho^T(x^{(j)}) dx^{(j)}. \end{aligned}$$

Note that $x^{(j)}$ is a vector in $\mathbb{R}^{|\Gamma(j)|}$. Interchanging the sum in the second term and let $\gamma = (\Gamma(1), \dots, \Gamma(|\pi|)) \leq \pi$, we obtain

$$\begin{aligned} s_k &= \sum_{\gamma \leq 1} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \rho(x) dx - \sum_{\gamma \leq 1} \sum_{\gamma \leq \pi < 1} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \left[\prod_{j=1}^{|\pi|} \rho^T(x^{(j)}) dx^{(j)} \right] \\ &= \sum_{\gamma \leq 1} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \left[\rho(x) - \sum_{\pi: \gamma \leq \pi < 1} \prod_{j=1}^{|\pi|} \rho^T(x^{(j)}) dx^{(j)} \right], \text{ here } x = (x^{(1)}, \dots, x^{(|\pi|)}), \\ &= \sum_{\gamma \leq 1} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \left[\rho(x) - \sum_{\pi \in \Pi(|\gamma|): \pi < 1} \rho^T(x, \pi) \right] dx \\ &= \sum_{\gamma \leq 1} \int_{\mathbb{R}^{|\gamma|}} h^\gamma(x) \rho^T(x) dx. \end{aligned}$$

This completes the induction step and the proof of Lemma 15. □

Appendix C. An explicit computation for the two-point correlation function of P_∞ . First thanks to translation and reflection invariance one has $\rho(s, t) = \rho(0, t - s) = \rho(0, s - t)$, therefore it suffices to compute $\rho(0, t)$ for $t > 0$. This function was computed in an earlier paper [25] (which also contains many interesting statistics about the real roots); we choose provide the details for the reader's convenience.

Let $\gamma(t) = e^{-t^2/2}$ and $g(t) = \gamma(t)P_\infty(t)$. The zero distribution of P_∞ and g are the same, so it suffices to compute the two-point correlation function for the real zeros of g . The covariance matrix for $g(0), g(t), g'(0), g'(t)$ is a symmetric 4×4 matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where A, B, C, D are 2×2 matrices. It follows that the conditional distribution of $(g'(0), g'(t))$ given $g(0) = 0$ and $g(t) = 0$ is a centered bivariate Gaussian with covariance matrix $\Sigma = D - CA^{-1}B$. Since $\mathbb{E}[g(t)g(s)] = \gamma(t - s)$,

one has $C = B^T$ and

$$A = \begin{pmatrix} 1 & \gamma(t) \\ \gamma(t) & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -t\gamma(t) \\ t\gamma(t) & 0 \end{pmatrix},$$

$$D = \begin{pmatrix} 1 & (1-t^2)\gamma(t) \\ (1-t^2)\gamma(t) & 1 \end{pmatrix}.$$

Therefore via explicit computation (below $\gamma = \gamma(t) = e^{-t^2/2}$)

$$(C.1) \quad \Sigma = I + M, \quad M = \begin{pmatrix} x & y \\ y & x \end{pmatrix}, \quad x = \frac{-t^2\gamma^2}{1-\gamma^2}, \quad y = (1-t^2)\gamma - \frac{t^2\gamma^3}{1-\gamma^2}.$$

Note that $(g(0), g(t))$ is a centered bivariate Gaussian with covariant A , thus the density at $(0, 0)$ is $(2\pi\sqrt{1-\gamma^2})^{-1}$. It follows that

$$(C.2) \quad \rho(0, t) = \frac{1}{2\pi\sqrt{1-\gamma^2}} \mathbb{E}[|X| \cdot |Y|]$$

where (X, Y) have mean zero bivariate Gaussian distributions with covariant Σ .

Note that $|\alpha| = \frac{1}{\sqrt{2\pi}} \int_0^\infty (1 - e^{-\alpha^2 x^2/2}) x^{-3/2} dx$ (which follows from a simple rescaling of the integration variable and integration by parts). Using the Kac-Rice formula, it follows that

$$\mathbb{E}[|X| \cdot |Y|] = \frac{1}{2\pi} \int_0^\infty \int_0^\infty (f(u, v) - f(u, 0) - f(0, v) + f(0, 0)) u^{-3/2} v^{-3/2} du dv$$

$$f(u, v) = \mathbb{E}[e^{-(uX^2 + vY^2)/2}] = \left| \det \left(I + \Sigma \begin{pmatrix} u & 0 \\ 0 & v \end{pmatrix} \right) \right|^{-1/2}$$

$$= ((1+u)(1+v) + x(u+v+2uv) + uv(x^2 - y^2))^{-1/2}.$$

Redefine $A = x$ and $B = x^2 - y^2$, we have

$$f(u, v) = (1 + (A+1)v + (1 + A + v + 2Av + Bv)u)^{-1/2}$$

$$= (1 + (A+1)v)^{-1/2} (1 + M_v u)^{-1/2}$$

where $M_v = \frac{1+A+(1+2A+B)v}{1+(A+1)v}$. We now use the elementary identity

$$\int (1+\alpha)^{-1/2} \alpha^{-3/2} d\alpha = -2(1+\alpha)^{1/2} \alpha^{-1/2} + C$$

and obtain, via a simple change of variables,

$$\begin{aligned}
 & \int_0^\infty (f(u, v) - f(0, v)) u^{-3/2} du \\
 &= \frac{1}{\sqrt{1 + (A+1)v}} \int_0^\infty ((1 + M_v u)^{-1/2} - 1) u^{-3/2} du \\
 &= \frac{1}{\sqrt{1 + (A+1)v}} (2u^{-1/2} - 2(1 + M_v u)^{1/2} u^{-1/2}) \Big|_0^\infty \\
 &= \frac{-2M_v^{1/2}}{\sqrt{1 + (A+1)v}}.
 \end{aligned}$$

In particular we could let $v = 0$ and obtain $-2(1 + (A+1)u)^{-1/2}$ as the result. Therefore

$$\mathbb{E}[|X| \cdot |Y|] = \frac{1}{2\pi} \int_0^\infty 2 \left(\sqrt{A+1} - \frac{(1 + A + (1 + 2A + B)v)^{1/2}}{1 + (A+1)v} \right) v^{-3/2} dv.$$

Let $N = \frac{1+2A+B}{(1+A)^2} < 1$ (since $B < A^2$). Using the change of variables $v \mapsto (1 + 2A + B)v/(1 + A)$ we obtain

$$\begin{aligned}
 \mathbb{E}[|X| \cdot |Y|] &= \frac{1}{2\pi} 2(1 + 2A + B)^{1/2} \int_0^\infty \left(1 - \frac{(1+v)^{1/2}}{1 + N^{-1}v} \right) v^{-3/2} dv \\
 &= \frac{1}{2\pi} 2(1 + 2A + B)^{1/2} \left(\frac{2}{\alpha} \arctan\left(\frac{u}{\alpha}\right) + \frac{2}{u} - 2v^{-1/2} \right) \Big|_{v=0}^{v=\infty}
 \end{aligned}$$

where $u := v^{1/2}(1+v)^{-1/2}$ and $\alpha = \sqrt{N/(1-N)}$. Thus

$$\mathbb{E}[|X| \cdot |Y|] = \frac{1}{2\pi} 4(1 + 2A + B)^{1/2} \left(\frac{1}{\alpha} \arctan\left(\frac{1}{\alpha}\right) + 1 \right).$$

Note that $\arctan(\frac{1}{\alpha}) = \arccos(\alpha/\sqrt{1+\alpha^2}) = \arccos(\sqrt{N})$. Therefore

$$\mathbb{E}[|X| \cdot |Y|] = \frac{1}{2\pi} \left(4(1 + 2A + B)^{1/2} + 4\sqrt{A^2 - B} \arccos\left(\frac{\sqrt{1+2A+B}}{1+A}\right) \right).$$

Recalling $A = x$ and $B = x^2 - y^2$ and (C.1) we obtain

$$(C.3) \quad \rho(0, t) = \frac{1}{\pi^2 \sqrt{1 - e^{-t^2}}} \left(\sqrt{(1+x)^2 - y^2} + |y| \arcsin\left(\frac{|y|}{1+x}\right) \right)$$

where x, y are defined using (C.1). One could check that $1+x \geq 0 \geq y$ thus by letting

$$\delta := \frac{|y|}{1+x} = \frac{e^{-t^2/2}(e^{-t^2/2} + t^2 - 1)}{1 - e^{-t^2} - t^2 e^{-t^2}}$$

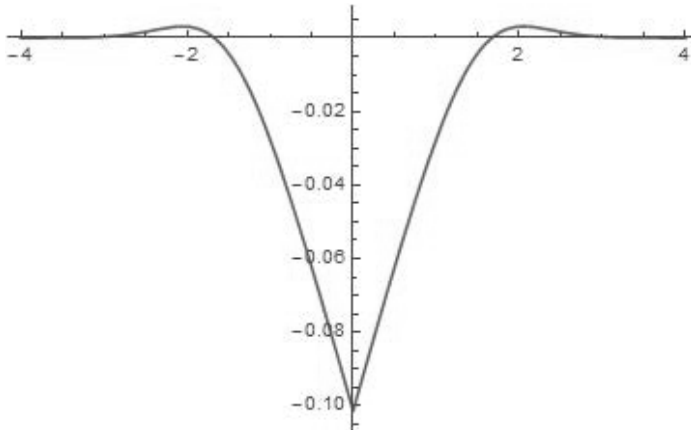


Figure 1. Mathematica plot of $k(t) := \rho(0, t) - \frac{1}{\pi^2}$.

we obtain

$$\rho(0, t) = \frac{\sqrt{(1 - e^{-t^2})^2 - t^4 e^{-t^2}}}{\pi^2 (1 - e^{-t^2})} \left(1 + \frac{\delta}{\sqrt{1 - \delta^2}} \arcsin \delta \right)$$

which recovers [25, (D8,D9)]. From here a numerical evaluation gives

COROLLARY 2. Let $k(t) = \rho(0, t) - \frac{1}{\pi^2}$. Then $\widehat{k}(0) + \frac{1}{\pi} = 0.18198 \dots > 0$.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA 22904-4137

E-mail: yendo@virginia.edu

DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, LEVEL 4, BLOCK S17, 10 LOWER KENT RIDGE ROAD, SINGAPORE 119076

E-mail: matvhv@nus.edu.sg

REFERENCES

- [1] R. B. Ash, *Probability and Measure Theory. With Contributions by Catherine Doléans-Dade*, 2nd ed., Harcourt/Academic Press, Burlington, MA, 2000.
- [2] J.-M. Azaïs, F. Dalmao, and J. R. León, CLT for the zeros of classical random trigonometric polynomials, *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016), no. 2, 804–820.
- [3] J.-M. Azaïs and J. R. León, CLT for crossings of random trigonometric polynomials, *Electron. J. Probab.* **18** (2013), no. 68, 17.

- [4] A. T. Bharucha-Reid and M. Sambandham, *Random Polynomials*, *Probab. Math. Statist.*, Academic Press, Orlando, FL, 1986.
- [5] F. Dalmao, Asymptotic variance and CLT for the number of zeros of Kostlan Shub Smale random polynomials, *C. R. Math. Acad. Sci. Paris* **353** (2015), no. 12, 1141–1145.
- [6] Y. Do, H. Nguyen, and V. Vu, Real roots of random polynomials: expectation and repulsion, *Proc. Lond. Math. Soc.* (3) **111** (2015), no. 6, 1231–1260.
- [7] Y. Do, O. Nguyen, and V. Vu, Roots of random polynomials with coefficients of polynomial growth, *Ann. Probab.* **46** (2018), no. 5, 2407–2494.
- [8] A. Edelman and E. Kostlan, How many zeros of a random polynomial are real?, *Bull. Amer. Math. Soc. (N.S.)* **32** (1995), no. 1, 1–37.
- [9] K. Farahmand, *Topics in Random Polynomials*, *Pitman Res. Notes Math. Ser.*, vol. 393, Longman, Harlow, 1998.
- [10] F. Götze, D. Kaliada, and D. Zaporozhets, Correlation functions of real zeros of random polynomials, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **454** (2016), 102–111.
- [11] A. Granville and I. Wigman, The distribution of the zeros of random trigonometric polynomials, *Amer. J. Math.* **133** (2011), no. 2, 295–357.
- [12] J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág, *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, *Univ. Lecture Ser.*, vol. 51, American Mathematical Society, Providence, RI, 2009.
- [13] I. A. Ibragimov and N. B. Maslova, The average number of zeros of random polynomials, *Vestnik Leningrad. Univ.* **23** (1968), no. 19, 171–172.
- [14] S. Janson, Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs, *Ann. Probab.* **16** (1988), no. 1, 305–312.
- [15] M. Kac, On the average number of real roots of a random algebraic equation, *Bull. Amer. Math. Soc.* **49** (1943), 314–320.
- [16] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation (III), *Rec. Math. [Mat. Sbornik]* **12(54)** (1943), 277–286.
- [17] ———, On the distribution of the zeros and α -values of a random integral function (I), *J. London Math. Soc.* **20** (1945), 130–136.
- [18] ———, On the distribution of zeros and α -values of a random integral function (II), *Ann. of Math.* (2) **49** (1948), 885–952.
- [19] N. B. Maslova, On the distribution of the number of real roots of random polynomials, *Theory Probab. Appl.* **19** (1974), 461–473.
- [20] ———, On the variance of the number of real roots of random polynomials, *Theory Probab. Appl.* **19** (1974), 35–52.
- [21] M. L. Mehta, *Random Matrices*, 2nd ed., Academic Press, Boston, MA, 1991.
- [22] F. Nazarov and M. Sodin, Correlation functions for random complex zeroes: strong clustering and local universality, *Comm. Math. Phys.* **310** (2012), no. 1, 75–98.
- [23] H. Nguyen, O. Nguyen, and V. Vu, On the number of real roots of random polynomials, *Commun. Contemp. Math.* **18** (2016), no. 4, 1550052, 17.
- [24] A. C. Schaeffer, Inequalities of A. Markoff and S. Bernstein for polynomials and related functions, *Bull. Amer. Math. Soc.* **47** (1941), 565–579.
- [25] G. Schehr and S. N. Majumdar, Real roots of random polynomials and zero crossing properties of diffusion equation, *J. Stat. Phys.* **132** (2008), no. 2, 235–273.
- [26] A. N. Shiryaev, *Probability. Translated from the first (1980) Russian edition by R. P. Boas*, 2nd ed., *Grad. Texts in Math.*, vol. 95, Springer-Verlag, New York, 1996.
- [27] Z. Su and Q. Shao, Asymptotics of the variance of the number of real roots of random trigonometric polynomials, *Sci. China Math.* **55** (2012), no. 11, 2347–2366.
- [28] G. Szegő, Über einen Satz von A. Markoff, *Math. Z.* **23** (1925), no. 1, 45–61.
- [29] T. Tao and V. Vu, Local universality of zeroes of random polynomials, *Int. Math. Res. Not. IMRN* **2015** (2015), no. 13, 5053–5139.
- [30] J. E. Wilkins Jr., An asymptotic expansion for the expected number of real zeros of a random polynomial, *Proc. Amer. Math. Soc.* **103** (1988), no. 4, 1249–1258.