HOLE RADII FOR THE KAC POLYNOMIALS AND DERIVATIVES

HOI H. NGUYEN AND OANH NGUYEN

Abstract. The Kac polynomial

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

with independent coefficients of variance 1 is one of the most studied models of random polynomials.

It is well-known that the empirical measure of the roots converges to the uniform measure on the unit disk. On the other hand, at any point on the unit disk, there is a hole in which there are no roots, with high probability. In a beautiful work [13], Michelen showed that the holes at ± 1 are of order $1/n$. We show that in fact, all the hole radii are of the same order. The same phenomenon is established for the derivatives of the Kac polynomial as well.

1. Introduction

Approximation by roots of polynomials of coefficients $\{-1, 0, 1\}$ is a classical and interesting topic in analysis, with fascinating pictures and conjectures. For instance, it follows from a result of Borwein and Pinner [3, Theorem 1] that for any given ζ of d-th root of unity, the distance from it to any $z \neq \zeta$ from the set of zeros of $\{-1, 0, 1\}$ polynomials of degree n which vanish at ζ of order at most k (i.e. $f^{(k+1)}(\zeta) \neq 0$) can be bounded by

$$|z - \zeta| \geq e^{-1} \frac{(k!)^{\lceil \phi(d)/2 \rceil}}{(n+1)^{(k+1)\lceil \phi(d)/2 \rceil + 1}},$$

where $\phi(d)$ is the usual Euler phi-function. This result is asymptotically optimal. On the other hand, the smallest distance can be sub-exponentially small as $n \to \infty$ if ζ is on the unit circle and not a root of unity (such as when ζ is an algebraic number of small Mahler measure), see for instance [3, Corollary 1, Theorem 3]. The situation at 1 is also interesting, it was shown from the same paper [3, Corollary 4, Theorem 6] that $|1 - z| \geq \frac{1}{n^{k+2}}$ (which is again near optimal) for any real roots z of $\{-1, 0, 1\}$ polynomials of degree n which vanish at 1 of order exactly k. Note that the distance is significantly larger if z is purely complex. See Figure 1. We also refer the reader to [3, 5, 19] and the references therein for further interesting discussions and problems.

Our goal in this note is to study the distances from some probabilistic viewpoint. More generally, consider the Kac polynomial

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

H.N. is supported by NSF CAREER grant DMS 1752345 and by the Simons Foundation, O.N. is supported by NSF grants DMS–1954174 and DMS–2246575. This work was initiated under the SQuaREs 2021 program of AIM, we thank the Institute for the generous support.

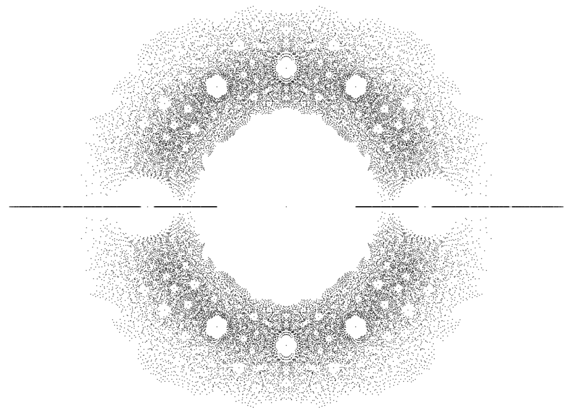

FIGURE 1. Zeros of all polynomials with ± 1 coefficients and degree at most eight ([3]).

where ξ_i are independent (not necessarily identically distributed), real-valued random variables with mean 0 and variance 1.

For this random polynomial, it is well-known that the empirical distribution of the roots converges to the uniform distribution on the unit circle ([9]). So, the roots concentrate near the unit circle. And in particular, the real roots concentrate near ± 1. However, precisely at ± 1, there are holes that do not contain any roots. It was conjectured by Shepp and Vanderbei [20], and confirmed recently by Michelen [13, Theorem 1.2] that the typical distance of *real roots* to 1 for random Kac polynomial is of order $O(1/n)$.

Theorem 1.1. *Let ξ_i be iid with mean zero and variance one. For any constant $\delta > 0$, there exists a constant $C > 0$ so that*

$$\mathbf{P}(\textit{there exists a real root in } [1 - C/n, 1 + C/n]) \geq 1 - \delta$$

for all n sufficiently large.

It is not hard to establish the lower bound and conclude that the hole at 1 (and -1) has radius of order $\Theta(1/n)$.

How about the hole radius at other points on the unit circle? A recent result by Cook, Yakir, Zeitouni and the first author [6] (see also Michelen and Sahasrabudhe [14] for the Gaussian case) shows that the distance between the zero set of $f_{n,0}$ and the unit circle is of order $\frac{1}{n^2}$. So, this is a lower bound for all hole radii.

From the result in [3] and Figure 1, it is natural to predict that the hole radii exhibit different orders at different points. For instance, in Figure 2 (source [2]) where all roots of polynomials with coefficients ± 1 and degree at most 24, one can observe that there are largest holes at ± 1, smaller holes possibly at the roots of unity, and barely visible holes at other points. In Figure 3, we draw sampled roots of random Kac polynomials with ± 1 coefficients and degree $n = 1000$. Note that as $n = 1000$ is large compared to 24, the holes are no longer visible in the figure unless being zoomed in properly. However, the striking similarities between the two figures would suggest that the same observation would remain true for large n. In Figure 4, we display sampled roots of the first derivative polynomial with ± 1 coefficients and degree $n = 1000$ which has the same pattern as in Figure 3.

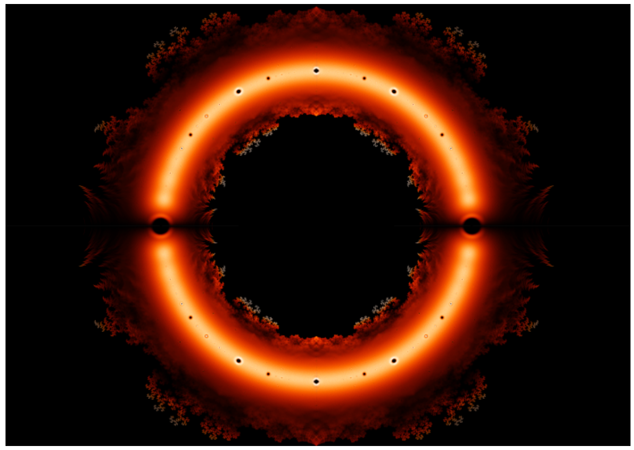

FIGURE 2. All roots of polynomials with coefficients ± 1 and degree at most 24 (source [2])

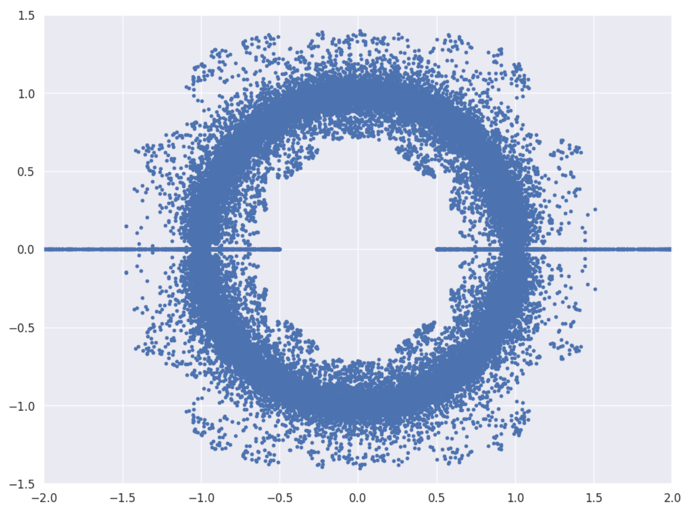

FIGURE 3. Sampled roots of the Kac polynomials with Rademacher coefficients

Disproving this prediction, in this paper, we show that for every point ζ on the unit circle, the precise order of the hole at ζ is $\Theta(1/n)$ (although the implied constant might depend on ζ). Moreover, we show that this holds also for the derivatives of the Kac polynomial. For a positive integer constant ρ, let us define

$$f_{\rho,n} = \sum_{i=0}^{n} a_{i,\rho,n}\xi_i z^i$$

where $a_{i,\rho,n} = (1 + o_n(1))i(i-1)\dots(i-\rho+1) \in \mathbb{R}$ with the $o_n(1)$ converges to 0 uniformly in i, as $n \to \infty$ and ρ fixed. For any set $S \subset \mathbb{C}$, let $N_{f_{\rho,n}}(S)$ be the number of roots of $f_{\rho,n}$ in S. If $\rho = 0$, we get the Kac polynomial. The ρ-th derivative of the Kac polynomial corresponds to $x^{-\rho} f_{\rho,n}(x)$. Here is our main result.

Theorem 1.2. *Let ρ be any positive integer constant. Assume that the random variables ξ_i are independent (and not necessarily identically distributed) with mean 0, variance 1 and bounded $(2 + \varepsilon_0)$-moment for some $\varepsilon_0 > 0$. For every $\zeta \in S^1$, the radius of the hole at ζ is*

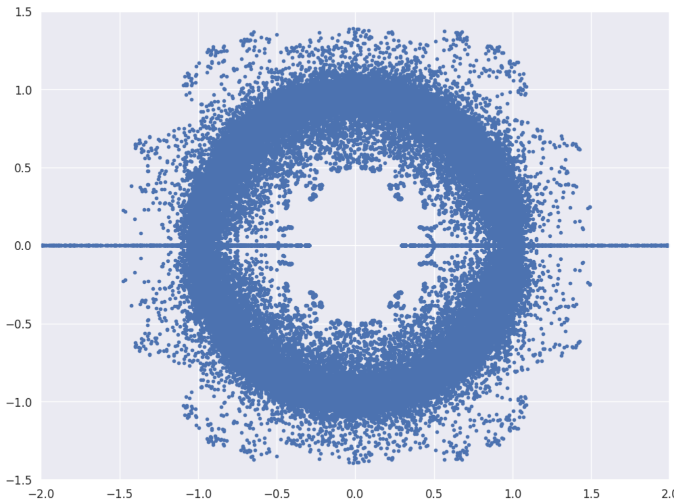

FIGURE 4. Sampled roots of the first derivative of Kac polynomials with Rademacher coefficients

$\Omega(1/n)$. *In particular, for every* $\varepsilon > 0$, *there exist positive constants* c_ρ *and* C_ρ *such that for all* $\zeta \in S^1$

$$\text{[Upper bound]} \qquad \mathbf{P}(N_{f_{\rho,n}}(B(\zeta, C_\rho/n)) > 0) \geq 1 - \varepsilon \tag{1}$$

and

$$\text{[Lower bound]} \qquad \mathbf{P}(N_{f_{\rho,n}}(B(\zeta, c_\rho/n)) = 0) \geq 1 - \varepsilon \tag{2}$$

for sufficiently large n.

As far as we understand, the proof in [13] is restricted to real roots and cannot be applied to complex roots.

For a somewhat related discussion, we refer to [4], [10, Section 5.1.3] and the references therein, where the authors studied the probability there is no zero in a given region.

2. Proof sketch and ingredients

For the upper bound, we need to split into two cases: $\zeta = \pm 1$ and $\zeta \neq \pm 1$. For the former, we show that the proof in [13] can be adapted to cover $f_{\rho,n}$ for general $\rho \in \mathbb{N}$. This method relies on the simple observation that if a polynomial f changes sign in an interval on the real line then it has at least one root there. Since this observation only holds for real roots, it merely works for $\zeta = \pm 1$ where one can reduce the upper-bound problem to showing that there exists a real root in the interval centered at ζ and radius C/n. For $\zeta \neq \pm 1$, we need a different approach.

To this end, we note that the expected value of $N_{f_{\rho,n}}(B(\zeta, C/n))$ being large does not imply that the number of real roots is non-zero with high probability. However, it can be achieved via Chebyshev inequality if we can show that its variance is of smaller order than its mean squared. To do so, the high-level idea is to show that $f_{\rho,n}$, after rescaled properly, converges to a Gaussian process, say f_∞ which we can estimate the growth of $\mathbf{Var}\left(N_{f_\infty}(B(\zeta, C/n))\right)$, the variance of the number of roots of f_∞ in the ball, in terms of C and then pass the result back to $f_{\rho,n}$. So, we consider a rescaled version of $f_{\rho,n}$ by zooming in at the local neighborhood of ζ as follows

$$g_n(z) = \frac{1}{n^{\rho+1/2}} f_{\rho,n}\left(\zeta + \frac{1}{n}\zeta z\right).$$

When the random variables are Gaussian, we know that g_n is a Gaussian process with covariance

$$\begin{aligned} \mathbf{E} g_n(z)\overline{g_n(w)} &= \frac{1}{n^{2\rho+1}} \sum_{k=0}^{n} |\zeta|^{2k} a_{k,\rho,n}^2 \left(1 + \frac{1}{n} z\right)^k \left(1 + \frac{1}{n}\bar{w}\right)^k \\ &= \frac{1}{n^{2\rho+1}} \sum_{k=0}^{n} a_{k,\rho,n}^2 \left(1 + \frac{1}{n} z\right)^k \left(1 + \frac{1}{n}\bar{w}\right)^k. \end{aligned}$$

We note that ζ disappears on the right-most side and could potentially account for why the hole radii are of the same order.

To show the cancellation $\mathbf{Var}\left(N_{f_\infty}(B(\zeta, C/n))\right) = o\left(\mathbf{E}\left(N_{f_\infty}(B(\zeta, C/n))\right)\right)^2$, we note that the variance is an integral of its correlation functions, namely an integral of $\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)$ which is in turn of order $|z_1 - z_2|^{-1}$. So, when z_1 and z_2 are far away, the integrand is small, accounting for the cancellation. This suggests that the numbers of roots in far-away regions are weakly correlated which is consistent with previously established results for real roots ([17]) and radius of complex roots ([6, 15, 16]). To handle the diagonal region when z_1 is near z_2, we come up with a simple argument, though via rather long and tedious algebraic manipulations, showing that the function $\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)$ is indeed continuous everywhere and hence the diagonal has negligible contribution, see Lemma 4.5. To carry out this strategy, we actually replace the whole ball $B(\zeta, C/n)$ by a subset, denoted by U_C, which is a thin strip along the unit circle. This is a major device that allows us to reduce from $|z_1 - z_2|$ to $|\mathrm{Im}(z_1) - \mathrm{Im}(z_2)|$ which reduces the dimension and facilitates the rather elegant proof that follows.

To pass from f_∞ to $f_{\rho,n}$, we need to show some sort of uniform integrability of $N_{f_\infty}(B(\zeta, C/n)$. For that, we adapt a double Taylor expansion argument used in [12] (see Lemma 4.6).

Finally, to establish the lower bound (2), we will show that the expected number of roots in the ball $B(\zeta, c/n)$ is small for sufficiently small c and then apply Markov's inequality. The derivation of the expected number of roots is first reduced to the Gaussian setting when all coefficients ξ_i are iid standard Gaussian, via the *universality* properties of the random polynomials. To do the calculation for the Gaussian case, there are two possible ways. The first way is to directly apply the classical Kac-Rice formula to $f_{\rho,n}$. The second way, which is what we perform here, is to derive it through f_∞ using the limits that we already establish for the upper bound.

Notations. For the rest of the paper, to simplify the notation, we will often drop the subscript ρ. For instance, we write f_n in place of $f_{\rho,n}$. For a function f, let $\mathcal{Z}(f) = \{z \in \mathbb{C} : f(z) = 0\}$ be the zero set of f.

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

If $a_n \ll b_n \ll a_n$, we say that $b_n = \Theta(a_n)$. If $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$, we say that $a_n = o(b_n)$. We also write that $a_n = O_C(b_n)$ if the implied constant depends on a given parameter C.

3. Proof of Theorem 1.2: upper bound for $\zeta = \pm 1$

When $\zeta = \pm 1$, Michelen [13] already showed the stated upper bound for the Kac polynomial $f_{0,n}$. We will show that this proof can be easily adapted to cover the general case $f_{\rho,n}$. We assume that $\zeta = 1$ as the case $\zeta = -1$ is completely similar. It suffices to show that with probability $\geq 1-\varepsilon$, there is at least one root of $f_{\rho,n}$ in the interval $J := [1-C/n, 1+C/n]$ for some large constant C. Let $f_{0,n}$ be the ρ-th anti-derivative of $z^{-\rho} f_{\rho,n}$, then

$$f_{0,n} = \sum_{i=0}^{n} (1+o(1))\xi_i z^i$$

is basically the Kac polynomial (if disregarding the $1+o(1)$ terms). By interlacing, this can be deduced from showing that there are at least $\rho+1$ roots of $f_{0,n}$ in the same interval. To this end, we show that we can find $(\rho+1)$ sub-intervals of J each of which observes a sign change of $f_{0,n}$ and hence contains at least one root.

Consider the rescaled polynomial

$$h_n(x) = \frac{1}{\sqrt{n}} f_{0,n}(1+x/n), \quad x \in \mathbb{R}.$$

Let M be a large constant and $x_1, \dots, x_M$ be deterministic points in J. By [13, Lemma 5] (which is a rather direct application of the Lindeberg-Feller Central Limit Theorem), the random vector $(h_n(x_1), \dots, h_n(x_M))$ converges to the Gaussian vector $(h(x_1), \dots, h(x_M))$ where h is a centered, real Gaussian process with covariance

$$\mathbf{E} h(x)h(y) = \int_0^1 e^{(x+y)t} dt.$$

(In fact, in [13], this result is established for the Kac polynomial without the $1+o(1)$ terms as above but the proof can easily go through without changes when these terms are present.)

Let $M = (\rho+1)K$ where K is a large constant to be chosen. By [13, Lemma 6], there exists a constant $\gamma > 0$ such that any centered Gaussian vector $(Z_1, \dots, Z_K)$ with variances $\mathbf{E} Z_i^2 = 1$ for all i and covariances $|\mathbf{E} Z_i Z_j| \leq \gamma$ for all $i \neq j$ satisfies

$$\mathbf{P}(\text{all } Z_1, \dots, Z_K \text{ have the same sign}) \leq 2^{-K+2}.$$

Direct calculation shows that if $y = \alpha x$ with $x > 1$ and $\alpha > 1$ then

$$\left| \mathbf{E} \frac{h(x)h(y)}{\sqrt{\mathbf{Var} h(x) \mathbf{Var} h(y)}} \right| = o_{\alpha\to\infty}(1).$$

So, for a given K, by taking α sufficiently large, we can make this number smaller than γ. We then take $x_i = \alpha^{i-1}$, $i = 1, \dots, M$. So, for $j = 0, \dots, \rho$,

$$\mathbf{P}(\text{all } h(x_{jK+1}), \dots, h(x_{jK+K}) \text{ have the same sign}) \leq 2^{-K+2}.$$

And so, for sufficiently large n,

$$\mathbf{P}(h_n \text{ does not have any real roots in } [1 + x_{jK+1}/n, 1 + x_{jK+K}/n]) \leq 2^{-K+3}.$$

By the union bound, the probability that h_n has less than $\rho+1$ real roots in $[1+x_1/n, 1+x_M/n]$ (which is a union of $\rho+1$ such intervals above) is at most $(\rho+1)2^{-K+3}$. By choosing K sufficiently large so that this number is smaller than ε, we obtain the desired tail probability.

4. Proof of Theorem 1.2: upper bound for $\zeta \neq \pm 1$

We want to show that there exists a constant C such that with probability at least $1-\varepsilon$, there is at least one root of $f_{\rho,n}$ in the ball $B(\zeta, 2C/n)$. For a sufficiently small constant $\delta > 0$ depending only on ρ, we consider the strip $\zeta + \zeta U_C/n$ that goes along the unit circle where $U_C = (-\delta, \delta) \times (-C, C) \subset B(0, 2C)$. Since this strip is a subset of $B(\zeta, 2C/n)$, it suffices to show that with probability at least $1-\varepsilon$, there is at least one root of $f_{\rho,n}$ in U_C. The use of U_C in place of the ball allows us to derive the upper bound using much simpler arguments because for $z_1, z_2 \in U_C$, they are either very close or $|z_1 - z_2| \approx |y_1 - y_2|$ where y_i is the imaginary part of z_i and is a real number!

Since the upper bound at ± 1 has been proved in Theorem 1.1, it suffices to assume that $\zeta \neq \pm 1$.

4.1. The setup.

Consider the following rescaled version of f_n, centered around ζ

$$g_n(z) = \frac{1}{n^{\rho+1/2}} f_{\rho,n}\left(\zeta + \frac{1}{n}\zeta z\right), z \in U_C. \tag{3}$$

The proof consists of the following steps.

(1) Construct a Gaussian process g_∞ that shall be the limit of g_n.

(2) Show that for all $\varepsilon \geq 0$, there exists C such that

$$\mathbf{P}(N_{g_\infty}(U_C) = 0) \leq \varepsilon. \tag{4}$$

(3) When the random variables ξ_i are iid standard Gaussian, show that on U_C, g_n weakly convergences to g (see Subsection 4.5)

$$g_n \xrightarrow{w} g. \tag{5}$$

(4) Show that for each positive integer k,

$$\mathbf{E}N_{g_n}^k(U_C) \to \mathbf{E}N_{g_\infty}^k(U_C) \tag{6}$$

for general ξ_i (not necessarily Gaussian).

(5) Show that this implies

$$\mathbf{P}(N_{g_n}(U_C) = 0) \to \mathbf{P}(N_{g_\infty}(U_C) = 0). \tag{7}$$

These steps are carried out in Sections 4.2, 4.3, 4.5, 4.6, 4.7, respectively.

4.2. Construct g_∞.

We have for all $z, w \in U_C$,

$$\begin{aligned}
\mathbf{E}g_n(z)\overline{g_n(w)} &= \frac{1}{n^{\rho+1}}\sum_{k=0}^{n}|\zeta|^{2k}a_{k,\rho,n}^2(1+\frac{1}{n}z)^k(1+\frac{1}{n}\bar{w})^k \\
&= (1+o_n(1))\frac{1}{n}(1+\frac{1}{n}z)^\rho(1+\frac{1}{n}\bar{w})^\rho\frac{\partial^{2\rho}}{\partial z^\rho\partial\bar{w}^\rho}\sum_{k=0}^{n}(1+\frac{1}{n}z)^k(1+\frac{1}{n}\bar{w})^k \\
&= (1+o_n(1))\frac{\partial^{2\rho}}{\partial z^\rho\partial\bar{w}^\rho}\frac{(1+\frac{1}{n}z)^{n+1}(1+\frac{1}{n}\bar{w})^{n+1}-1}{n\left((1+\frac{1}{n}z)(1+\frac{1}{n}\bar{w})-1\right)}.
\end{aligned}$$

where we used $|\zeta| = 1$. When $\rho = 0$, we have

$$\mathbf{E} g_n(z)\overline{g_n(w)} \xrightarrow{n\to\infty} \frac{1}{z+\bar{w}}(\exp(z+\bar{w})-1) = \int_0^1 e^{(z+\bar{w})t}dt = F(z+\bar{w}).$$

where

$$F(u) := \int_0^1 e^{tu}dt = \frac{e^u - 1}{u}. \tag{8}$$

Similarly, for all ρ, we have

$$\mathbf{E} g_n(z)\overline{g_n(w)} \xrightarrow{n\to\infty} \frac{\partial^{2\rho}}{\partial z^\rho \partial \bar{w}^\rho} F(z+\bar{w}) = F^{(2\rho)}(z+\bar{w}) = \int_0^1 t^{2\rho} e^{t(z+\bar{w})}dt.$$

Moreover, since $\zeta \neq \pm 1$,

$$\mathbf{E} g_n(z) g_n(w) = (1+o(1))\frac{\partial^{2\rho}}{\partial z^\rho \partial w^\rho} \frac{\zeta^{2n+2}(1+\frac{1}{n}z)^{n+1}(1+\frac{1}{n}w)^{n+1}-1}{n\left(\zeta^2(1+\frac{1}{n}z)(1+\frac{1}{n}w)-1\right)} \xrightarrow{n\to\infty} 0$$

as the denominator blows up (with its derivatives bounded) and the numerator is bounded. So, it is logical to define the tentative limit g_∞ to be

$$g_\infty(z) = \int_0^1 t^\rho e^{zt} dB(t)$$

where B is the standard complex Brownian motion $B(t) = \frac{1}{\sqrt{2}}(B_1(t) + \sqrt{-1}B_2(t))$ with B_1, B_2 being independent standard real Brownian motions. We shall prove in Section 4.5 that g_n indeed converges to g_∞ when $\zeta \neq \pm 1$.

4.3. **Upper bound the hole radius for g_∞.** In this section, we want to show that there exists at least one root of g_∞ in U_C with high probability (as $C \to \infty$). By Chebyshev's inequality, it suffices to show the following

$$\mathbf{Var} N_{g_\infty}(U_C) = o_{C\to\infty}(\mathbf{E} N_{g_\infty}(U_C))^2. \tag{9}$$

Indeed, we have

$$\mathbf{P}(N_{g_\infty}(U_C) = 0) \le \frac{\mathbf{Var} N_{g_\infty}(U_C)}{(\mathbf{E} N_{g_\infty}(U_C))^2} = o(1).$$

So, by choosing C to be sufficiently large, the probability of g_∞ having no roots can be arbitrarily small.

To prove (9), let us evaluate $\mathbf{E} N_{g_\infty}(U_C)$. Since g_∞ is a Gaussian analytic function ([10]), we can use Kac-Rice formula for Gaussian case. By [10, Corollary 3.4.2], we have

$$\mathbf{E} N_{g_\infty}(U_C) = \int_{U_C} \rho_1(z) := \int_{U_C} \frac{V - T^2 S^{-1}}{\pi S} dz = \int_{U_C} \frac{VS - T^2}{\pi S^2} dz$$

where

$$S(u) = \mathbf{E}|g_\infty(z)|^2, \quad T(u) = \mathbf{E} g'_\infty(z)\overline{g_\infty(z)}, \quad V(u) = \mathbf{E}|g'_\infty(z)|^2.$$

Let $u = z + \bar{z} \in \mathbb{R}$. By the definition (8) of F, we have

$$S = F^{(2\rho)}(u).$$

Taking derivative gives

$$T = F^{(2\rho+1)}(u), V = F^{(2\rho+2)}(u).$$

Since

$$F(u) = \frac{e^u - 1}{u} = \sum_{k=0}^{\infty} \frac{u^k}{(k+1)!},$$

we have

$$F^{(k)}(0) = \frac{1}{k+1}.$$

Therefore,

$$F^{(\rho+2)}(0)F^{(\rho)}(0) - \left(F^{(\rho+1)}(0)\right)^2 \neq 0. \tag{10}$$

And so, $V(0)S(0) - T^2(0) = \frac{1}{(2\rho+1)(2\rho+3)} - \frac{1}{(2\rho+2)^2} > 0$. By choosing δ to be sufficiently small, it holds that for all $|u| \leq 2\delta$, we have

$$V(u)S(u) - T^2(u) = \Theta_\rho(1) \quad \text{and} \quad S(u) = \Theta_\rho(1). \tag{11}$$

By the definition of U_C, $u = z + \bar{z} \in [-2\delta, 2\delta]$ for all $z \in U_C$. Hence, $\rho_1(z) = \Theta_\rho(1)$ which implies

$$\mathbf{E}N_{g_\infty}(U_C) = \Theta(C).$$

It remains to show the following.

Lemma 4.1. *We have*

$$\mathbf{Var}N_{g_\infty}(U_C) \ll C \log C. \tag{12}$$

Proof of Lemma 4.1. Let $N = N_{g_\infty}(U_C)$. We have

$$\mathbf{Var}(N_{g_\infty}(U_C)) = \mathbf{E}N(N-1) + \mathbf{E}N = \mathbf{E}N(N-1) + O(C).$$

So, it suffices to show that $\mathbf{E}N(N-1) \ll C\log C$. By [10, Corollary 3.4.2],

$$\mathbf{E}N(N-1) = \int_{U_C}\int_{U_C} \rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2) dz_1 dz_2 \tag{13}$$

where ρ_2 is the 2-point correlation function which can be derived using the following formula (also [10, Corollary 3.4.2])

$$\pi^2 \rho_2(z_1, z_2) = \frac{\mathrm{per}(\mathcal{V} - \mathcal{T}\mathcal{S}^{-1}\mathcal{T}^*)}{\det(\mathcal{S})}, \tag{14}$$

with $\mathcal{S}, \mathcal{T}, \mathcal{V}$ being 2×2 matrices defined by

$$\mathcal{S}_{i,j} = \mathbf{E}g(z_i)\bar{g}(z_j) = \int_0^1 t^{2\rho} e^{tu_{ij}} dt = F^{(2\rho)}(u_{ij}), \quad u_{ij} = z_i + \bar{z}_j$$

$$\mathcal{T}_{ij} = \mathbf{E}g'(z_i)\bar{g}(z_j) = \int_0^1 t^{2\rho+1} e^{tu_{ij}} dt = F^{(2\rho+1)}(u_{ij})$$

$$V_{ij} = \mathbf{E}g'(z_i)\overline{g'}(z_j) = \int_0^1 t^{2\rho+2} e^{tu_{ij}} dt = F^{(2\rho+2)}(u_{ij}).$$

Since $\mathcal{T}^* = \mathcal{T}$, by letting $\mathcal{W} = \mathcal{T}\mathcal{S}^{-1}\mathcal{T}^*$, we have

$$\mathcal{W}_{ij} = = \sum_{k=1}^{2}\sum_{\ell=1}^{2} \mathcal{T}_{ik}(\mathcal{S}^{-1})_{k\ell}\mathcal{T}_{\ell j} \tag{15}$$

where

$$\det(\mathcal{S})\mathcal{S}^{-1} = \begin{pmatrix} \mathcal{S}_{22} & -\mathcal{S}_{12} \\ -\mathcal{S}_{21} & \mathcal{S}_{11} \end{pmatrix}.$$

We present a straightforward observation from the forms of ρ_1 and ρ_2 that if $\mathcal{S}, \mathcal{T}, \mathcal{V}, \mathcal{W}$ were diagonal (namely, setting the off-diagonal entries to 0), then $\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2) = 0$. Motivated by this, we will show that for z_1 and z_2 far away, the above matrices are indeed diagonally dominated, and hence $\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)$ is small. In particular, we show the following.

Lemma 4.2 (Off diagonal)**.** *For all D_0 satisfying $10 \le D_0 \le C$, let $\mathcal{D}_{C,D_0} = \{(z_1, z_2) \in U_C^2 : |z_1 - z_2| \ge D_0\}$. For all $(z_1, z_2) \in U_C$, it holds that*

$$|\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)| \ll |z_1 - z_2|^{-1}.$$

In particular, we have

$$\iint_{\mathcal{D}_{C,D_0}} \rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2) dz_1 dz_2 \ll C \log(C/D_0).$$

When z_1 and z_2 are close, we show that $\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)$ is bounded and hence the contribution from the diagonal region is negligible.

Lemma 4.3 (Diagonal)**.** *There exists a constant M, independent of C, such that for all $(z_1, z_2) \in U_C^2$, we have*

$$|\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)| \ll M.$$

This implies

$$\iint_{U_C^2 \setminus \mathcal{D}_{C,D_0}} \left(\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)\right) dz_1 dz_2 \ll C D_0.$$

Assuming these lemmas, letting $D_0 = \log C$, we get that $\mathbf{Var} N(U_C) \ll C \log C$ as desired. This finishes the proof of Lemma 4.1. $\square$

Proof of Lemma 4.2. We write $x_i = \mathrm{Re}(z_i)$ and $y_i = \mathrm{Im}(z_i)$ for $i = 1, 2$. Note that $|x_i| \le \delta$ for all i. By (11),

$$|\mathcal{S}_{11}| = \Theta(1), |\mathcal{S}_{22}| = \Theta(1). \tag{16}$$

Let $|z_1 - z_2| = D$, we have $D \ge D_0 \ge 10$. So, $|y_1 - y_2| \gg D$ and hence

$$|u_{12}| = |u_{21}| \gg D. \tag{17}$$

For all $u \in \mathbb{C}$, since $(uF(u))^{(k)} = kF^{(k-1)} + uF^{(k)}$ and since the left-hand side equals e^u for all k, we get

$$F^{(k)}(u) = \frac{e^u - kF^{(k-1)}(u)}{u}.$$

For $u_{12} = z_1 + \bar{z}_2$, we have $|\mathrm{Re}(u_{12})| \le 2\delta$ and so,

$$|e^{u_{12}}| = e^{\mathrm{Re}(u_{12})} = O(1).$$

Hence, by (17) and induction in k, it holds for all $k \le \rho$ that

$$|F^{(k)}(u_{12})| \ll \frac{1}{|u_{12}|}$$

which gives

$$|\mathcal{S}_{12}| = O(D^{-1}). \tag{18}$$

Similarly, $|\mathcal{S}_{21}| = O(D^{-1})$.

Thus,

$$\det(\mathcal{S}) = \mathcal{S}_{11}\mathcal{S}_{22} - O(D^{-2}) = (1 + O(D^{-2}))\mathcal{S}_{11}\mathcal{S}_{22} = \Theta(1).$$

And so,

$$\mathcal{S}^{-1} = \begin{pmatrix} \frac{1+O(D^{-2})}{\mathcal{S}_{11}} & O(D^{-1}) \\ O(D^{-1}) & \frac{1+O(D^{-2})}{\mathcal{S}_{22}} \end{pmatrix}.$$

Similarly, the same bounds as in (16) and (18) hold for $\mathcal{T}$ and $\mathcal{V}$ in place of $\mathcal{S}$.

Using these bounds, we get that $\mathcal{W}_{12}$ is the sum of 4 terms each of which is of order $O(D^{-1})$. So, $|\mathcal{W}_{12}| = O(D^{-1})$. Likewise, $|W_{21}| = O(D^{-1})$.

Similarly,

$$\mathcal{W}_{11} = O(D^{-1}) + \mathcal{T}_{11}^2(\mathcal{S}^{-1})_{11} = O(D^{-1}) + \frac{(1 + O(D^{-2}))\mathcal{T}_{11}^2}{\mathcal{S}_{11}} = \frac{\mathcal{T}_{11}^2}{\mathcal{S}_{11}} + O(D^{-1}) = \Theta(1)$$

and

$$\mathcal{W}_{22} = \frac{\mathcal{T}_{22}^2}{\mathcal{S}_{22}} + O(D^{-1}) = \Theta(1).$$

Therefore,

$$|(\mathcal{V} - \mathcal{W})_{12}| = O(D^{-1}), |(\mathcal{V} - \mathcal{W})_{12}| = O(D^{-1}).$$

And

$$(\mathcal{V} - \mathcal{W})_{11} = \mathcal{V}_{11} - \frac{\mathcal{T}_{11}^2}{\mathcal{S}_{11}} + O(D^{-1}) = \Theta(1)$$

where in the last equality, we used (11). So,

$$\operatorname{per}(\mathcal{V} - \mathcal{W}) = \left(\mathcal{V}_{11} - \frac{\mathcal{T}_{11}^2}{\mathcal{S}_{11}}\right)\left(\mathcal{V}_{22} - \frac{\mathcal{T}_{22}^2}{\mathcal{S}_{22}}\right) + O(D^{-1}).$$

All in all, we get

$$\begin{aligned} \pi^2 \left|\rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2)\right| &= \frac{\left(V_{11} - \frac{T_{11}^2}{\mathcal{S}_{11}}\right)\left(V_{22} - \frac{T_{22}^2}{\mathcal{S}_{22}}\right) + O(D^{-1})}{(1 + O(D^{-2}))\mathcal{S}_{11}\mathcal{S}_{22}} - \frac{\left(V_{11} - \frac{T_{11}^2}{\mathcal{S}_{11}}\right)\left(V_{22} - \frac{T_{22}^2}{\mathcal{S}_{22}}\right)}{\mathcal{S}_{11}\mathcal{S}_{22}} \\ &= \frac{O(D^{-1})}{\Theta(1)} = O(D^{-1}). \end{aligned}$$

Integrating this over $\mathcal{D}_{C,D_0}$ we get

$$\begin{aligned} \iint_{\mathcal{D}_{C,D_0}} \rho_2(z_1, z_2) - \rho_1(z_1)\rho_1(z_2) dz_1 dz_2 &\ll \iint_{\mathcal{D}_{C,D_0}} |z_1 - z_2|^{-1} dz_1 dz_2 \\ &\ll \int_{U_C} dz_1 \int_{w \in [-2\delta, 2\delta 2] \times [-2C, 2C], |w| \ge D_0} |w|^{-1} dw \\ &= \Theta(C \int_{D_0}^{C} y^{-1} dy) = \Theta(C \log(C/D_0)). \end{aligned}$$

This finishes the proof of Lemma 4.2. □

Proof of Lemma 4.3. Since $\rho_1(z_1)\rho_1(z_2)$ is bounded over U_C^2, we only need to show the boundedness of ρ_2. By the first part of Lemma 4.2, we can reduce to the region

$$U_{\text{diag}} := \{(z_1, z_2) \in U_C^2 : |z_1 - z_2| \ll 1\}.$$

Since ρ_2 can be written as a function of $(u_{ij})_{i,j=1,2}$, it is also a function of x_1, x_2 and $\Delta_y := y_1 - y_2$ where we recall $x_i = \mathrm{Re}(z_i)$ and $y_i = \mathrm{Im}(z_i)$. Note that U_{diag} is a subset of $\{(z_1, z_2) : |x_1| \le \delta, |x_2| \le \delta, \Delta_y \ll 1\}$ which is a compact set. So, if we can show that ρ_2 is in fact a continuous function of x_1, x_2 and $\Delta_y := y_1 - y_2$, we conclude that it is bounded U_{diag}. To show continuity, note that the only possible singularities of ρ_2 occur when $\det(\mathcal{S}) = 0$. Thus, it suffices to show the following.

Lemma 4.4. *If* $\det(\mathcal{S}(z_1, z_2)) = 0$ *then* $z_1 = z_2$.

Lemma 4.5. *For all* $z \in U_C$, $\rho_2(z_1, z_2)$ *is continuous at* $(z_1, z_2) = (z, z)$.

Proving these lemmas will complete the proof of Lemma 4.3. □

Proof of Lemma 4.4. Assume that $\det(\mathcal{S}(z_1, z_2)) = 0$. Since $\mathcal{S}(z_1, z_2)$ is a 2×2 (complex) matrix, there exist deterministic complex numbers w_1, w_2 such that

$$(w_1\ w_2)\mathcal{S}(w_1\ w_2)^{\mathcal{T}} = 0.$$

In other words,

$$\mathbf{E}|w_1 g(z_1) + w_2 g(z_2)|^2 = 0.$$

Since the left-hand side equals $\int_0^1 |w_1 e^{tz_1} + w_2 e^{tz_2}|^2 dt$, we conclude that the integrand is 0 for almost all t (and hence for all t by continuity). Therefore, it is necessary that $z_1 = z_2$. □

Next we give a direct proof for Lemma 4.5, where we note that there might be other ways to justify it by using the methods of [1, 15, 16].

Proof of Lemma 4.5. We need to show that for all $z \in \mathbb{C}$ (or just U_C if needed),

$$\lim_{(\varepsilon,\delta)\to(0,0)} \rho_2(z, z + \varepsilon + i\delta) \text{ exists.}$$

We shall perform Taylor expansion to the order 2 of the functions appearing in (14). Here $z_1 = z = x + iy, z_2 = z + \varepsilon + i\delta$. Then

$$u_{11} = 2x =: u, u_{12} = (2x + \varepsilon) - i\delta, u_{22} = 2x + 2\varepsilon.$$

Let

$$a = F^{(2\rho)}(u), b = F^{(2\rho+1)}(u), c = F^{(2\rho+2)}(u), d = F^{(2\rho+3)}(u), e = F^{(2\rho+4)}(u).$$

So,

$$\mathcal{S}_{11} = F^{(2\rho)}(u) = a,$$

$$\mathcal{S}_{22} = F^{(2\rho)}(u + 2\varepsilon) \approx a + 2\varepsilon b + 2\varepsilon^2 c,$$

$$\mathcal{S}_{12} = \overline{\mathcal{S}_{21}} = F^{(2\rho)}(u+\varepsilon-i\delta) \approx a+(\varepsilon-i\delta)b+\frac{(\varepsilon-i\delta)^2}{2}c = \left(a + \varepsilon b + \frac{\varepsilon^2-\delta^2}{2}c\right) - i\left(\delta b + \varepsilon\delta c\right).$$

That is

$$\mathcal{S} \approx \begin{pmatrix} a & \left(a + \varepsilon b + \frac{\varepsilon^2-\delta^2}{2}c\right) - i\left(\delta b + \varepsilon\delta c\right) \\ \left(a + \varepsilon b + \frac{\varepsilon^2-\delta^2}{2}c\right) + i\left(\delta b + \varepsilon\delta c\right) & a + 2\varepsilon b + 2\varepsilon^2 c \end{pmatrix}. \tag{19}$$

The denominator of ρ_2 is

$$\begin{aligned}\det(\mathcal{S}) &= \mathcal{S}_{11}\mathcal{S}_{22} - \mathcal{S}_{12}\mathcal{S}_{21} \approx a(a + 2\varepsilon b + 2\varepsilon^2 c) - \left(a + \varepsilon b + \frac{\varepsilon^2-\delta^2}{2}c\right)^2 - (\delta b + \varepsilon\delta c)^2 \\ &= 2\varepsilon^2 ac - \varepsilon^2 b^2 - (\varepsilon^2 - \delta^2)ac - \delta^2 b^2 + o(\varepsilon^2 + \delta^2) \approx (\varepsilon^2 + \delta^2)(ac - b^2).\end{aligned}$$

Note that by (11), $ac - b^2 \neq 0$.

Since $\mathcal{T}$ and $\mathcal{V}$ are similar to $\mathcal{S}$, we get

$$\mathcal{T} \approx \begin{pmatrix} b & \left(b + \varepsilon c + \frac{\varepsilon^2-\delta^2}{2}d\right) - i\left(\delta c + \varepsilon\delta d\right) \\ \left(b + \varepsilon c + \frac{\varepsilon^2-\delta^2}{2}d\right) + i\left(\delta c + \varepsilon\delta d\right) & b + 2\varepsilon c + 2\varepsilon^2 d \end{pmatrix}$$

and

$$\mathcal{V} \approx \begin{pmatrix} c & \left(c + \varepsilon d + \frac{\varepsilon^2-\delta^2}{2}e\right) - i\left(\delta d + \varepsilon\delta e\right) \\ \left(c + \varepsilon d + \frac{\varepsilon^2-\delta^2}{2}e\right) + i\left(\delta d + \varepsilon\delta e\right) & c + 2\varepsilon d + 2\varepsilon^2 e \end{pmatrix}.$$

Note that the numerator of (14) for ρ_2 equals

$$\operatorname{per}(\mathcal{V} - \mathcal{W}) = \frac{1}{\det(\mathcal{S})^2}\operatorname{per}((\det\mathcal{S})\mathcal{V} - (\det\mathcal{S})\mathcal{W}).$$

So far, (14) becomes

$$\pi^2\rho_2 \approx \frac{\operatorname{per}((\det\mathcal{S})\mathcal{V} - (\det\mathcal{S})\mathcal{W})}{(\varepsilon^2+\delta^2)^3(ac-b^2)^3} =: \frac{\operatorname{per}(\mathcal{Y})}{(\varepsilon^2+\delta^2)^3(ac-b^2)^3}.$$

Next, we write down $(\det\mathcal{S})\mathcal{W}$. We have

$$(\det\mathcal{S})\mathcal{W}_{11} = \mathcal{T}_{11}\mathcal{S}_{22}\mathcal{T}_{11} - \mathcal{T}_{12}\mathcal{S}_{21}\mathcal{T}_{11} - \mathcal{T}_{11}\mathcal{S}_{12}\mathcal{T}_{21} + \mathcal{T}_{12}\mathcal{S}_{11}\mathcal{T}_{21}$$
$$(\det\mathcal{S})\mathcal{W}_{12} = \mathcal{T}_{11}\mathcal{S}_{22}\mathcal{T}_{12} - \mathcal{T}_{12}\mathcal{S}_{21}\mathcal{T}_{12} - \mathcal{T}_{11}\mathcal{S}_{12}\mathcal{T}_{22} + \mathcal{T}_{12}\mathcal{S}_{11}\mathcal{T}_{22}$$
$$(\det\mathcal{S})\mathcal{W}_{21} = \mathcal{T}_{21}\mathcal{S}_{22}\mathcal{T}_{11} - \mathcal{T}_{22}\mathcal{S}_{21}\mathcal{T}_{11} - \mathcal{T}_{21}\mathcal{S}_{12}\mathcal{T}_{21} + \mathcal{T}_{22}\mathcal{S}_{11}\mathcal{T}_{21}$$
$$(\det\mathcal{S})\mathcal{W}_{22} = \mathcal{T}_{21}\mathcal{S}_{22}\mathcal{T}_{12} - \mathcal{T}_{22}\mathcal{S}_{21}\mathcal{T}_{12} - \mathcal{T}_{21}\mathcal{S}_{12}\mathcal{T}_{22} + \mathcal{T}_{22}\mathcal{S}_{11}\mathcal{T}_{22}$$

So,

$$\begin{aligned}
\mathcal{Y}_{11} &= (\varepsilon^2+\delta^2)(ac-b^2)c - (\mathcal{T}_{11}\mathcal{S}_{22}\mathcal{T}_{11} - \mathcal{T}_{12}\mathcal{S}_{21}\mathcal{T}_{11} - \mathcal{T}_{11}\mathcal{S}_{12}\mathcal{T}_{21} + \mathcal{T}_{12}\mathcal{S}_{11}\mathcal{T}_{21}) \\
&= (\varepsilon^2+\delta^2)(ac-b^2)c - b^2(a + 2\varepsilon b + 2\varepsilon^2 c) \\
&\quad +2b\operatorname{Re}\left(\left(\left(b + \varepsilon c + \frac{\varepsilon^2-\delta^2}{2}d\right) + i\left(\delta c + \varepsilon\delta d\right)\right)\left(\left(a + \varepsilon b + \frac{\varepsilon^2-\delta^2}{2}c\right) - i\left(\delta b + \varepsilon\delta c\right)\right)\right) \\
&\quad -a\left(b + \varepsilon c + \frac{\varepsilon^2-\delta^2}{2}d\right)^2 - a\left(\delta c + \varepsilon\delta d\right)^2 \\
&= (\varepsilon^2+\delta^2)(ac-b^2)c - b^2(a + 2\varepsilon b + 2\varepsilon^2 c) \\
&\quad +2b\left(\left(b + \varepsilon c + \frac{\varepsilon^2-\delta^2}{2}d\right)\left(a + \varepsilon b + \frac{\varepsilon^2-\delta^2}{2}c\right) + \left(\delta c + \varepsilon\delta d\right)\left(\delta b + \varepsilon\delta c\right)\right) \\
&\quad -a\left(b + \varepsilon c + \frac{\varepsilon^2-\delta^2}{2}d\right)^2 - a\left(\delta c + \varepsilon\delta d\right)^2.
\end{aligned}$$

Grouping the terms with ε, δ, $\varepsilon\delta$, ε^2, δ^2 and smaller order terms together, we get

$$\begin{aligned}
\mathcal{Y}_{11} &= \varepsilon\left(-2b^3 + 2b^3 + 2abc - 2abc\right) + \varepsilon^2\left((ac-b^2)c + b^2c + abd - ac^2 - abd\right) \\
&\quad +\delta^2\left((ac-b^2)c - b^2c - abd + 2b^2c + abd - ac^2\right) + O(\varepsilon^3 + \varepsilon\delta^2 + \varepsilon^2\delta + \delta^3) \\
&= O(\varepsilon^3 + \varepsilon\delta^2 + \varepsilon^2\delta + \delta^3).
\end{aligned}$$

We now try to accomplish the same estimate for the other three entries of $\mathcal{Y}$. We have

$$\begin{aligned}
\mathcal{Y}_{22} &= (\varepsilon^2+\delta^2)(ac-b^2)\left(c+2\varepsilon d+2\varepsilon^2 e\right) - (\mathcal{T}_{22}\mathcal{S}_{11}\mathcal{T}_{22} - \mathcal{T}_{22}\mathcal{S}_{21}\mathcal{T}_{12} - \mathcal{T}_{21}\mathcal{S}_{12}\mathcal{T}_{22} + \mathcal{T}_{21}\mathcal{S}_{22}\mathcal{T}_{12}) \\
&= (\varepsilon^2+\delta^2)(ac-b^2)\left(c+2\varepsilon d+2\varepsilon^2 e\right) - a\left(b+2\varepsilon c+2\varepsilon^2 d\right)^2 \\
&\quad +2\left(b+2\varepsilon c+2\varepsilon^2 d\right)\mathrm{Re}\left(\left(\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)+i\left(\delta c+\varepsilon\delta d\right)\right)\right. \\
&\qquad\qquad \left.\left(\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c\right)-i\left(\delta b+\varepsilon\delta c\right)\right)\right] \\
&\quad -\left(b+2\varepsilon c+2\varepsilon^2 d\right)\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)^2 - \left(b+2\varepsilon c+2\varepsilon^2 d\right)\left(\delta c+\varepsilon\delta d\right)^2.
\end{aligned}$$

So,

$$\begin{aligned}
\mathcal{Y}_{22} &= (\varepsilon^2+\delta^2)(ac-b^2)\left(c+2\varepsilon d+2\varepsilon^2 e\right) - a\left(b+2\varepsilon c+2\varepsilon^2 d\right)^2 \\
&\quad +2\left(b+2\varepsilon c+2\varepsilon^2 d\right)\left(\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c\right)+\delta^2\left(c+\varepsilon d\right)\left(b+\varepsilon c\right)\right) \\
&\quad -\left(a+2\varepsilon b+2\varepsilon^2 c\right)\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)^2 - \left(a+2\varepsilon b+2\varepsilon^2 c\right)\delta^2\left(c+\varepsilon d\right)^2
\end{aligned}$$

Comparing this with $\mathcal{Y}_{11}$, we get

$$\begin{aligned}
\mathcal{Y}_{22} &= \mathcal{Y}_{11} + (\varepsilon^2+\delta^2)(ac-b^2)\left(2\varepsilon d+2\varepsilon^2 e\right) + (2\varepsilon b^3+2\varepsilon^2 b^2 c-4\varepsilon^2 ac^2-4\varepsilon^2 abd-4\varepsilon abc) \\
&\quad +2\left(2\varepsilon c+2\varepsilon^2 d\right)\left(\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c\right)+\left(\delta c+\varepsilon\delta d\right)\left(\delta b+\varepsilon\delta c\right)\right) \\
&\quad -\left(2\varepsilon b+2\varepsilon^2 c\right)\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)^2 - \left(2\varepsilon b+2\varepsilon^2 c\right)\delta^2\left(c+\varepsilon d\right)^2 + O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3) \\
&= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3) + (2\varepsilon b^3+2\varepsilon^2 b^2 c-4\varepsilon^2 ac^2-4\varepsilon^2 abd-4\varepsilon abc) \\
&\quad +2\left(2\varepsilon c+2\varepsilon^2 d\right)\left(b+\varepsilon c\right)\left(a+\varepsilon b\right) - \left(2\varepsilon b+2\varepsilon^2 c\right)\left(b+\varepsilon c\right)^2
\end{aligned}$$

which gives

$$\begin{aligned}
\mathcal{Y}_{22} &= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3) + (2\varepsilon b^3+2\varepsilon^2 b^2 c-4\varepsilon^2 ac^2-4\varepsilon^2 abd-4\varepsilon abc) \\
&\quad +2\left(2\varepsilon c+2\varepsilon^2 d\right)\left(ab+\varepsilon b^2+\varepsilon ac\right) - \left(2\varepsilon b+2\varepsilon^2 c\right)\left(b^2+2\varepsilon bc\right) \\
&= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3) + (2\varepsilon b^3+2\varepsilon^2 b^2 c-4\varepsilon^2 ac^2-4\varepsilon^2 abd-4\varepsilon abc) \\
&\quad +\left(4\varepsilon abc+4\varepsilon^2 b^2 c+4\varepsilon^2 ac^2+4\varepsilon^2 abd\right) - \left(2\varepsilon b^3+6\varepsilon^2 b^2 c\right) \\
&= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3).
\end{aligned}$$

Finally,

$$
\begin{aligned}
\mathcal{Y}_{12} &= (\varepsilon^2+\delta^2)(ac-b^2)\left(\left(c+\varepsilon d+\frac{\varepsilon^2-\delta^2}{2}e\right)-i\left(\delta d+\varepsilon\delta e\right)\right) \\
&\quad -(T_{11}\mathcal{S}_{22}T_{12}-T_{12}\mathcal{S}_{21}T_{12}-T_{11}\mathcal{S}_{12}T_{22}+T_{12}\mathcal{S}_{11}T_{22}) \\
&= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3)+(\varepsilon^2+\delta^2)(ac-b^2)c \\
&\quad -b\left(a+2\varepsilon b+2\varepsilon^2 c\right)\left(\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)-i\left(\delta c+\varepsilon\delta d\right)\right) \\
&\quad +\left(\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)-i\left(\delta c+\varepsilon\delta d\right)\right)^2\left(\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c\right)+i\left(\delta b+\varepsilon\delta c\right)\right) \\
&\quad +b\left(b+2\varepsilon c+2\varepsilon^2 d\right)\left(\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c\right)-i\left(\delta b+\varepsilon\delta c\right)\right) \\
&\quad -a\left(b+2\varepsilon c+2\varepsilon^2 d\right)\left(\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d\right)-i\left(\delta c+\varepsilon\delta d\right)\right).
\end{aligned}
$$

So,

$$
\begin{aligned}
\mathcal{Y}_{12} &= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3)+(\varepsilon^2+\delta^2)(ac-b^2)c \\
&\quad -ab\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d-i\left(\delta c+\varepsilon\delta d\right)\right)-2\varepsilon b^2(b+\varepsilon c-i\delta c)-2\varepsilon^2 b^2 c \\
&\quad +b\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d-i\left(\delta c+\varepsilon\delta d\right)\right)\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c+i\left(\delta b+\varepsilon\delta c\right)\right) \\
&\quad +\varepsilon c\left((b+\varepsilon c)-i\delta c\right)\left(a+\varepsilon b+i\delta b\right)+\frac{\varepsilon^2-\delta^2}{2}dab \\
&\quad -i\delta c\left((b+\varepsilon c)-i\delta c\right)\left(a+\varepsilon b+i\delta b\right)-i\varepsilon\delta abd \\
&\quad +b^2\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c-i\left(\delta b+\varepsilon\delta c\right)\right)+2\varepsilon bc\left(a+\varepsilon b-i\delta b\right)+2\varepsilon^2 abd \\
&\quad -ab\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d-i\left(\delta c+\varepsilon\delta d\right)\right)-2\varepsilon ac\left(b+\varepsilon c-i\delta c\right)+2\varepsilon^2 abd
\end{aligned}
$$

giving

$$
\begin{aligned}
\mathcal{Y}_{12} &= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3)+(\varepsilon^2+\delta^2)(ac-b^2)c\\
&\quad -ab\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d-i\left(\delta c+\varepsilon\delta d\right)\right)-2\varepsilon b^2(b+\varepsilon c-i\delta c)-2\varepsilon^2b^2c\\
&\quad +b^2\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c+i\left(\delta b+\varepsilon\delta c\right)\right)+\varepsilon bc\left(a+\varepsilon b+i\delta b\right)+\frac{\varepsilon^2-\delta^2}{2}abd\\
&\quad -i\delta bc\left(a+\varepsilon b+i\delta b\right)-i\varepsilon\delta abd\\
&\quad +\varepsilon bc\left(a+\varepsilon b+i\delta b\right)+\varepsilon^2c^2\left(a+\varepsilon b+i\delta b\right)-i\varepsilon\delta ac^2+\frac{\varepsilon^2-\delta^2}{2}dab\\
&\quad -i\delta bc\left(a+\varepsilon b+i\delta b\right)-i\varepsilon\delta ac^2-\delta^2ac^2-i\varepsilon\delta abd\\
&\quad +b^2\left(a+\varepsilon b+\frac{\varepsilon^2-\delta^2}{2}c-i\left(\delta b+\varepsilon\delta c\right)\right)+2\varepsilon bc\left(a+\varepsilon b-i\delta b\right)+2\varepsilon^2abd\\
&\quad -ab\left(b+\varepsilon c+\frac{\varepsilon^2-\delta^2}{2}d-i\left(\delta c+\varepsilon\delta d\right)\right)-2\varepsilon ac\left(b+\varepsilon c-i\delta c\right)+2\varepsilon^2abd\\
&= O(\varepsilon^3+\varepsilon\delta^2+\varepsilon^2\delta+\delta^3).
\end{aligned}
$$

Since the product of any two terms in $\{\varepsilon^3,\varepsilon\delta^2,\varepsilon^2\delta,\delta^3\}$ is bounded by $(\varepsilon^2+\delta^2)^3$, we yield the continuity of ρ_2. $\square$

4.4. **More on $N_{g_n}(U_C)$ and $N_{g_\infty}(U_C)$.** Before moving on the next section to show the convergence of g_n to g_∞ and their number of roots, we will first show that $N_{g_n}(U_C)$ have uniformly bounded higher moments.

Lemma 4.6. *There exists a constant $A=A(C)$ such that the following holds. For any $\ell\ge 0$, we have*

$$\mathbf{E}N_{g_n}^{\ell}(U_C)\le (A\ell)^{\ell}.$$

Proof. Let $k\ge A\ell$ for some large constant A. We want to bound the probability that $N_{g_n}(U_C)\ge k$. We divide U_C into $O(C\eta^{-2})$ (possibly overlapping) open balls $B_i=B(c_i,\eta)$ centered at c_i of radius η, which is chosen to be sufficiently small. Then there exists i such that B_i contains at least $s=k\eta^2/C$ roots. Then by Hermite interpolation, as g_n is analytic with probability one, we have

$$|g_n(c_i)|\le \frac{1}{s!}\eta^s\sup_{z\in B(c_i,\eta)}|g_n^{(s)}(z)|. \tag{20}$$

By Taylor expanding $g_n^s(z)$ around c_i, we obtain for any $m\ge 0$ (we later choose $m=\log n$),

$$|g_n^{(s)}(z)|\le \sum_{j=s}^{s+m-1}\frac{|g_n^{(j)}(c_i)|}{(j-s)!}\eta^{j-s}+\sup_{w\in B(c_i,\eta)}\frac{|g_n^{(s+m)}(w)|}{m!}\eta^m. \tag{21}$$

For each j, $g_n^{(j)}(c_i)$ is a Gaussian random variable with mean 0 and variance equals that of $\frac{1}{n^{j+\rho+1/2}}f_n^{(j)}(\zeta+\frac{1}{n}\zeta c_i)$, which is of order

$$(1+o(1))\frac{1}{n^{2j+2\rho+1}}\sum_{h=j}^{n}h^2(h-1)^2\dots(h-j+1)^2a_{h,\rho,n}^2|\zeta+\frac{1}{n}\zeta c_i|^{2h-2j}=O_C(1)$$

where we used $|\zeta + \frac{1}{n}\zeta c_i|^j \le (1+2C/n)^n \le e^{2C} = O_C(1)$.
So, by Gaussianity, for all $M_j \ge 1$,

$$\mathbf{P}(|g_n^{(j)}(c_i)| \gg M_j) \le e^{-M_j}.$$

Finally, for the supremum term, we observe

$$\begin{aligned}\sup_{w\in B(c_i,\eta)} |g_n^{(s+m)}(w)| &\ll \frac{1}{n^{s+m+\rho+1/2}} \sum_{h=s+m}^{n} h(h-1)\dots(h-s-m+1)|a_{h,\rho,n}||\xi_h|(|c_i|+\eta)^h \\ &\ll \frac{1}{n^{1/2}}\sum_{h=1}^{n}|\xi_h|.\end{aligned}$$

Note that if we hadn't used another round of Taylor expansion in (21) and just applied the above bound to $|g_n^{(s)}(z)|$ and take supremum, the term $n^{-1/2}\sum_{h=1}^n |\xi_h|$, which can be as large as $\sqrt{n}$, would be too big to handle. Here, we performed (21) so that the extra term $\eta^m/m!$ would swallow the $n^{-1/2}\sum_{h=1}^n |\xi_h|$. Indeed, for an M_0 to be chosen,

$$\mathbf{P}(\sum_{h=1}^{n}|\xi_h| \ge M_0 n) \le n\mathbf{P}(|\xi_h| \ge M_0) \le ne^{-M_0}.$$

Thus,

$$\mathbf{P}(\sup_{w\in B(c_i,\eta)} |g_n^{(s+m)}(w)| \gg n^{1/2}M_0) \ll ne^{-M_0}.$$

Combining all of these events, we conclude that with probability at least $1 - ne^{-M_0} - \sum_{j=s}^{s+m-1} e^{-M_j}$, we have

$$\sup_{z\in B(c_i,\eta)} |g_n^{(s)}(z)| \ll \sum_{j=s}^{s+m-1} \frac{M_j}{(j-s)!}\eta^{j-s} + \frac{n^{1/2}M_0\eta^m}{m!}.$$

On this event,

$$|g_n(c_i)| \ll \frac{1}{s!}\eta^s \left(\sum_{j=s}^{s+m-1} \frac{M_j}{(j-s)!}\eta^{j-s} + \frac{n^{1/2}M_0\eta^m}{m!}\right)$$

which only happens with probability at most

$$O_C(1)\frac{1}{s!}\eta^s \left(\sum_{j=s}^{s+m-1} \frac{M_j}{(j-s)!}\eta^{j-s} + \frac{n^{1/2}M_0\eta^m}{m!}\right)$$

since $g_n(c_i)$ is a Gaussian random variable with variance $\Theta_C(1)$. All together, we get that the probability that $N^{\ell}_{g_n}(U_C) \ge k$ is at most (up to a constant depending on C),

$$\frac{C}{\eta^2}\left[\frac{1}{s!}\eta^s \left(\sum_{j=s}^{s+m-1} \frac{M_j}{(j-s)!}\eta^{j-s} + \frac{n^{1/2}M_0\eta^m}{m!}\right) + \sum_{j=s}^{s+m-1} e^{-M_j}\right] + ne^{-M_0}$$

for any choice of η, M_0, M_j, with $s = k\eta^2/C$. For instance, we choose $s = 8\ell$, we get $\eta = \frac{\sqrt{8C\ell}}{\sqrt{k}} \le \sqrt{\frac{8C}{A}}$. By setting

$$M_j = s\log\frac{1}{\eta} + \eta^{-(j-s)/2}, \quad M_0 = 2\ell\log k + \log n,$$

we obtain the tail probability of

$$\frac{C}{\eta^2}\left[\frac{1}{s!}\eta^s\left(e^{\sqrt{\eta}}+\frac{n^{1/2}M_0\eta^m}{m!}\right)+\eta^s\right]+k^{-2\ell}.$$

Sending $m\to\infty$, the term with m goes to 0, so we end up with

$$\eta^{s-2}+k^{-2\ell}\ll (C\ell)^{4\ell}k^{-4\ell+1}+k^{-2\ell}.$$

So,

$$\begin{aligned}\mathbf{E}N_{g_n}^{\ell}(U_C) &\ll (A\ell)^{\ell}+\ell\sum_{k=A\ell}^{\infty}k^{\ell-1}\mathbf{P}(N_{g_n}(U_C)\geq k)\\ &\ll (A\ell)^{\ell}+\ell\sum_{k=A\ell}^{\infty}\left((C\ell)^{4\ell}k^{-3\ell}+k^{-\ell-1}\right)\ll (A\ell)^{\ell}\end{aligned}$$

as desired.

□

4.5. **Convergence of g_n to g_∞ when $\tilde{\xi}_i$ are iid $\mathcal{N}(0,1)$.** [1] Now, we prove (5). We first start with two simple results for the Gaussian models.

Lemma 4.7. *With probability one, g_n and g_∞ do not have double roots in U_C.*

Proof. For g_n, if it has a double root then f_n also has a double root. As this is a polynomial of degree n, if $f_n(z)$ and $f_n'(z)$ have common roots then the resultant must have zero determinant. But the resultant is a non-degenerate multivariate function of the Gaussians, so it is zero with probability zero.

For $g_\infty(z)$, for any $\alpha>0$, we divide U_C into $O(C\alpha^{-2})$ balls B_i of radius α. We will show that the probability there exists i such that N_i, the number of zeros in B_i, is greater than 2 is of order $O(\alpha^4)$, from which we see that the given probability will be bounded by $O(\alpha^2)$ after taking union bound. Indeed, using the boundedness of ρ_2 in Lemma 4.3,

$$\mathbf{P}(N_i\geq 2)\leq \mathbf{E}(N_i(N_i-1))=\int_{B_i\times B_i}\rho_2(z_1,z_2)dz_1dz_2\leq O(|B_i|\times|B_i|)=O(\alpha^4).$$

Sending α to 0, we conclude that the probability that g_∞ has double roots is 0. □

Our next simple result is the following.

Claim 4.8. *With probability one, $g_n(z)$ and g_∞ do not have roots on the boundary ∂U_C of U_C.*

Proof. We will show for g_∞ as the treatment for g_n is similar. From (11), we saw that for all $\alpha>0$ sufficiently small, $\rho_1(z)=O(1)$ for all $z\in U_C+B(0,\alpha)$. Let N be the number of roots in $\partial U_C+B(0,\alpha)$, then

$$\mathbf{P}(N\geq 1)\leq \mathbf{E}N=\int_{\partial U_C+B(0,\alpha)}\rho_1(z)dz\leq O_C(\alpha).$$

Sending α to 0, we obtain the claim. □

[1] We reserve the notation of ξ_i for random variables of general distribution (in the spirit of Theorem 1.2), while for Gaussian we use $\tilde{\xi}_i$.

Our treatment below is similar to [11, Section 4] where instead of real roots, we consider complex roots. First, let $\mathcal{H}$ be the set of all analytic function on the entire complex plane. We endow $\mathcal{H}$ with the topology of uniform convergence on the compact sets, which can be generated by the complete separable metric

$$d(f,g)=\sum_k \frac{1}{2^k}\frac{\|f-g\|_{\bar{D}_k}}{1+\|f-g\|_{\bar{D}_k}},$$

where $\bar{D}_k=\{z\in\mathbb{C}:|z|\le k\}$ and $\|f\|_K=\sup_{z\in K}|f(z)|$.

Lemma 4.9. *Let A_C be the set of all $f\in\mathcal{H}$ which do not have multiple roots in U_C and do not have roots over the boundary of U_C. Then the set A_C is open.*

Proof. This follows from Hurwitz's theorem. Indeed, consider a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$, which converges to some $f\in A_C$ locally uniformly. We will show that $f_n\in A_C$ for sufficiently large n. Let $R>0$ be large such that $U_C\subset D_R=\{z:|z|<R\}$. Let $z_1,\dots,z_d$ be the collections of all zeros of f in D_R with multiplicities $m_1,\dots,m_d$. Let $\alpha>0$ be sufficiently small such that the open disks z_i+D_α are disjoint, and do not intersect the boundary of the open sets D_R and of U_C, except when z_i are on one of these boundaries. By Hurwitz's theorem for sequence of (locally convergent) analytic functions, there exists n_0 such that for all $n\ge n_0$, f_n has exactly m_k zeros in z_k+D_α. Now if $z_i\in U_C$, then as $f\in A_C$, we must have $m_i=1$, and f_n has exactly one zero in z_i+D_α. Thus, $f_n\in A_C$ for all $n\ge n_0$. □

Lemma 4.10. *The mapping $f\to\mathcal{Z}_{U_C}(f)=\{z\in U_C:f(z)=0\}$ to the space of locally finite point measures on U_C endowed with the vague topology is continuous on A_C.*

Proof. This also follows from Hurwitz's theorem with the same argument as in the proof of Lemma 4.9, by letting the radius α tend to zero. □

Lemma 4.11. *We have the following weak convergence (of random elements with values in the metric space $\mathcal{H}$)*

$$g_n\xrightarrow{w}g_\infty.$$

Proof. By Prokhorov's theorem, it suffices to verify convergence in finite dimensional and tightness. Let $z_1,\dots,z_k$ be complex numbers. We first observe that the convergence in distribution of the Gaussian vector $(g_n(z_1),\dots,g_n(z_k))$ to the Gaussian vector $(g_\infty(z_1),\dots,g_\infty(z_k))$ already follows from our previous computations verifying the convergences of $\mathbf{E}g_n(z_i)\overline{g_n(z_j)}$ and $\mathbf{E}g_n(z_i)g_n(z_j)$ to $\mathbf{E}g_\infty(z_i)\overline{g_\infty(z_j)}$ and $\mathbf{E}g_\infty(z_i)g_\infty(z_j)$, respectively.

We need to verify tightness, for this, it suffices to show that for any $R>0$, there exists $C_R<\infty$ such that

$$\sup_n\sup_{|z|\le R}\mathbf{E}|g_n(z)|^2<C_R.$$

However, this is clear as

$$\mathbf{E}|g_n(z)|^2=(1+o_n(1))\frac{\partial^{2\rho}}{\partial z^\rho\partial\bar{w}^\rho}\frac{(1+\frac{1}{n}z)^{n+1}(1+\frac{1}{n}\bar{w})^{n+1}-1}{n\big((1+\frac{1}{n}z)(1+\frac{1}{n}\bar{w})-1\big)}\Big|_{w=z}.$$

□

Theorem 4.12. *We have that $N_{g_n}(U_C)\to N_{g_\infty}(U_C)$ in distribution and for each $k\in\mathbb{N}$, $\lim_{n\to\infty}\mathbf{E}N^k_{g_n}(U_C)=\mathbf{E}N^k_{g_\infty}(U_C)$.*

Proof. We have that $g_n \to g_\infty$ weakly, they are analytic and with probability one, they all belong to A_C. By Lemma 4.10, the point process $\mathcal{Z}_{U_C}(g_n)$ converges to $\mathcal{Z}_{U_C}(g_\infty)$ weakly, and hence the number of zeros $N_{g_n}(U_C)$ converges to $N_{g_\infty}(U_C)$ in distribution. In particular, for all $m \in \mathbb{N}$, $p_{n,m} := \mathbf{P}(N_{g_n}(U_C) = m) \to \mathbf{P}(N_{g_\infty}(U_C) = m) := p_m$ as $n \to \infty$. By Fatou's lemma and Lemma 4.6, it holds for all $\ell \in \mathbb{N}$ that

$$\mathbf{E}N_{g_\infty}^\ell(U_C) \le \liminf_n \mathbf{E}N_{g_n}^\ell(U_C) \le (A\ell)^\ell. \tag{22}$$

Fix $k \in \mathbb{N}$, we have for a large constant M,

$$\begin{aligned}
& \left|\mathbf{E}N_{g_n}^k(U_C) - \mathbf{E}N_{g_\infty}^k(U_C)\right| \\
\le\ & \sum_{m=0}^{M-1} m^k|p_{nm} - p_m| + \mathbf{E}N_{g_n}^k(U_C)\mathbf{1}_{N_{g_n}(U_C)\ge M} + \mathbf{E}N_{g_\infty}(U_C)\mathbf{1}_{N_{g_\infty}^k(U_C)\ge M} \\
\le\ & \sum_{m=0}^{M-1} m^k|p_{nm} - p_m| + 2\sup_{\hat{n}}\left(\mathbf{E}N_{g_{\hat{n}}}^{2k}(U_C)\right)^{1/2}\mathbf{P}\left(N_{g_{\hat{n}}}(U_C) \ge M\right)^{1/2} \text{ by Jensen's inequality} \\
\le\ & \sum_{m=0}^{M-1} m^k|p_{nm} - p_m| + 2(Ak)^k\sqrt{\frac{A}{M}} \text{ by (22) and Markov's inequality.}
\end{aligned}$$

Letting M and n go to infinity, we obtain the convergence in moments. $\square$

4.6. **Convergence for the number of real roots.** In this section, we prove (6). In other words, we prove the following convergence of the number of roots U_C. Note that the random variables are not necessarily Gaussian here.

The following generalizes Theorem 4.12 to non-Gaussian random variables.

Theorem 4.13. *Let C be a fixed positive number. For all $k \ge 0$, we have*

$$\mathbf{E}N_{g_n}^k(U_C) \to \mathbf{E}N_{g_\infty}^k(U_C)$$

as $n \to \infty$.

Let $\tilde{g}_n$ be the version of g_n when the random variables ξ_i are iid standard Gaussian. By Theorem 4.12, we have

$$\mathbf{E}N_{\tilde{g}_n}^k(U_C) \to \mathbf{E}N_{g_\infty}^k(U_C). \tag{23}$$

We note that the same proof holds with U_C replaced by $U_C + B(0,\alpha)$.

Proof. Note that the number of roots of g_n in U_C is the same as the number of roots in the original function f_n in the set $\zeta + \frac{1}{n}\zeta U_C$, by (3). For a small constant α, let φ be a test function approximating the indicator of $(U_C)^k$, in particular, we let φ be a smooth function such that

$$\mathbf{1}_{(U_C)^k} \le \varphi \le \mathbf{1}_{(U_C+B(0,\alpha))^k} \tag{24}$$

and $|\nabla^a\varphi(z)| \ll 1$ for all multi-indices a with $0 \le |a| \le 2k+4$.

By [7, Theorem 2.4] [2] applied to the function $G = \varphi$ and the centers $z_1 = \cdots = z_k = \zeta$, we get

$$\Bigg|\mathbf{E} \sum_{\zeta_{i_1},\dots,\zeta_{i_k} \in \mathcal{Z}(g_n)} \varphi(n(\zeta_{i_1}/\zeta - 1), \dots, n(\zeta_{i_1}/\zeta - 1))$$
$$-\mathbf{E} \sum_{\zeta_{i_1},\dots,\zeta_{i_k} \in \mathcal{Z}(\tilde{g}_n)} \varphi(n(\zeta_{i_1}/\zeta - 1), \dots, n(\zeta_{i_1}/\zeta - 1))\Bigg| \ll n^{-c}$$

where $c > 0$ is a small constant. Here, we note that the transformation $z := n(\zeta_i/\zeta - 1)$ is just the inverse of the rescaling map $\zeta_i = \zeta + \frac{1}{n}\zeta z$ that brings the neighborhood of ζ to U_C. We note that when φ is replaced by $\mathbf{1}_{(U_C)^k}$, the term under the expectation becomes $N^k(U_C)$. So, we have from (24) that

$$\mathbf{E}N^k_{g_n}(U_C) \le \mathbf{E}N^k_{\tilde{g}_n}(U_C + B(0, \alpha)) + O_\alpha(n^{-c}).$$

Using (23), we obtain

$$\limsup_{n\to\infty} \mathbf{E}N^k_{g_n}(U_C) \le \limsup_{n\to\infty} \mathbf{E}N^k_{\tilde{g}_n}(U_C + B(0, \alpha)) = \mathbf{E}N^k_{g_\infty}(U_C + B(0, \alpha)).$$

Sending α to 0, we obtain

$$\limsup_{n\to\infty} \mathbf{E}N^k_{g_n}(U_C) \le \limsup_{\alpha\to 0} \mathbf{E}N^k_{g_\infty}(U_C + B(0, \alpha)) = \mathbf{E}N^k_{g_\infty}(U_C)$$

where the last equality follows from the dominated convergence theorem, knowing that $\mathbf{E}N^k_{\tilde{g}_\infty}(U_C + B(0, \alpha)) < \infty$ for some $\alpha > 0$ (by (22)). Similarly, we get the reverse direction and conclude the proof. $\square$

4.7. Upper bound the hole radius for g_n.

In this section, we show the following theorem.

Theorem 4.14. *The random variables $N_{g_n}(U_C)$ converges to $N_{g_\infty}(U_C)$ in distribution as $n \to \infty$. In particular, we have (7):*

$$\mathbf{P}(N_{g_n}(U_C) = 0) \to \mathbf{P}(N_{g_\infty}(U_C) = 0).$$

Here, we recall that since the random variables $N_{g_n}(U_C)$ are discrete random variables supported on $\mathbb{N}$, convergence in distribution means convergence of the probability density $\mathbf{P}(N_{g_n}(U_C) = i)$, as i varies.

Proof. By Theorem 4.13, $N_{g_n(U_C)}$ converges to $N_{g_\infty}(U_C)$ in moments. By (22) and the Carleman's criteria (see [8]), $N_{g_\infty}(U_C)$ is uniquely determined by its moments. Thus, we infer that $N_{g_n(U_C)}$ converges to $N_{g_\infty}(U_C)$ in distribution. $\square$

[2] or perhaps a slightly readable [18, Theorem 4.3] which was written for the Kac polynomial but it holds also for the derivatives of the Kac polynomial.

5. Proof of Theorem 1.2: Lower bound

We want to show that for any $\varepsilon > 0$,there exists $c = c(\varepsilon, \zeta)$ such that

$$\mathbf{E}(N_{f_n}(B(\zeta, c/n))) \leq \varepsilon. \tag{25}$$

Without loss of generality, we assume that $\varepsilon < 1/100$ and $c < \varepsilon$.

The first step is to reduce to the Gaussian case, via universality results. Consider the Gaussian version of f_n,

$$\tilde{f}_n = \tilde{f}_{\rho,n} = \sum_{i=0}^{n} a_{i,\rho,n}\tilde{\xi}_i z^i$$

where $\tilde{\xi}_i$ are iid standard Gaussian.

Let G be a smooth function such that approximates the indicator of the ball, or more specifically, $\mathbf{1}_{B(\zeta,c/n)} \leq G \leq \mathbf{1}_{B(\zeta,2c/n)}$ and $||\nabla^a G||_\infty = O(n^a)$ for all $a \leq 3$. We now apply a universality property of f_n established in [7, Theorem 2.3]. This theorem applied to the function G states that the linear statistics $\mathbf{E}\sum_{w\in\mathcal{Z}(f_n)} G(w)$ is universal, i.e.,

$$\mathbf{E}\sum_{w\in\mathcal{Z}(f_n)} G(w) - \mathbf{E}\sum_{w\in\mathcal{Z}(\tilde{f}_n)} G(w) \ll n^{-\gamma}$$

for a constant γ independent of n and ζ.

Using this, we obtain

$$\begin{aligned}\mathbf{E}(N_{f_n}(B(\zeta, c/n))) &\leq \mathbf{E}\sum_{w\in\mathcal{Z}(f_n)} G(w) = \mathbf{E}\sum_{w\in\mathcal{Z}(\tilde{f}_n)} G(w) + o(1) \\ &\leq \mathbf{E}(N_{\tilde{f}_n}(B(\zeta, 2c/n))) + o(1).\end{aligned}$$

Thus, it suffices to prove that

$$\mathbf{E}(N_{\tilde{f}_n}(B(\zeta, 2c/n))) \leq \varepsilon/2. \tag{26}$$

In other words, it suffices to prove for the Gaussian case. To this end, we let $B_c = B(0, 2c)$ and define the functions g_n and g_∞ as before. We apply the Kac-Rice formula to g_∞ to get

$$\mathbf{E}N_{g_\infty}(B_c) = \int_{B_c} \rho_1(z).$$

By (11), we have for all $z \in B_c$, $\rho_1(z) \ll 1$. Thus,

$$\mathbf{E}N_{g_\infty}(B_c) \ll c^2 \leq \varepsilon/4.$$

By the same argument as for U_C (noting that $B_c \subset U_C$ for small c and for $C \geq c$), we obtain the same limit as in Theorem 4.12. So, we get

$$\lim_{n\to\infty} \mathbf{E}N_{g_n}(B_c) = \mathbf{E}N_{g_\infty}(B_c) \leq \varepsilon/4.$$

So, by choosing n to be sufficiently small, we obtain (26) as desired.

6. Acknowledgment

We thank Manjunath Krishnapur for suggesting helpful references on correlation functions.

References

[1] Michele Ancona and Thomas Letendre. Zeros of smooth stationary Gaussian processes. *Electron. J. Probab.*, 26:Paper No. 68, 81, 2021.

[2] John Baez, Dan Christensen, and Sam Derbyshire. The beauty of roots. *https://math.ucr.edu/home/baez/roots/beauty_web.pdf*.

[3] Peter Borwein and Christopher Pinner. Polynomials with $\{0, +1, -1\}$ coefficients and a root close to a given point. *Canad. J. Math.*, 49(5):887–915, 1997.

[4] Jeremiah Buckley, Alon Nishry, Ron Peled, and Mikhail Sodin. Hole probability for zeroes of gaussian taylor series with finite radii of convergence. *Probability Theory and Related Fields*, 171:377–430, 2018.

[5] Danny Calegari, Sarah Koch, and Alden Walker. Roots, Schottky semigroups, and a proof of Bandt's conjecture. *Ergodic Theory Dynam. Systems*, 37(8):2487–2555, 2017.

[6] Nicholas A Cook, Hoi H Nguyen, Oren Yakir, and Ofer Zeitouni. Universality of poisson limits for moduli of roots of kac polynomials. *International mathematics research notices*, 2023(8):6648–6690, 2023.

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

[8] Richard Durrett. Probability: Theory and examples.

[9] P. Erdös and P. Turán. On the distribution of roots of polynomials. *Ann. of Math. (2)*, 51:105–119, 1950.

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

[11] Alexander Iksanov, Zakhar Kabluchko, and Alexander Marynych. Local universality for real roots of random trigonometric polynomials. *Electron. J. Probab.*, 21:Paper No. 63, 19, 2016.

[12] Manjunath Krishnapur, Erik Lundberg, and Oanh Nguyen. The number of limit cycles bifurcating from a randomly perturbed center. *arXiv preprint arXiv:2112.05672*, 2021.

[13] Marcus Michelen. Real roots near the unit circle of random polynomials. *Transaction of AMS*, 2020.

[14] Marcus Michelen and Julian Sahasrabudhe. Random polynomials: the closest roots to the unit circle. *arXiv preprint arXiv:2010.10869*, 2020.

[15] Marcus Michelen and Oren Yakir. Fluctuations in the logarithmic energy for zeros of random polynomials on the sphere, 2023.

[16] Fedor Nazarov and Mikhail Sodin. Correlation functions for random complex zeroes: strong clustering and local universality. *Communications in Mathematical Physics*, 310(1):75–98, 2012.

[17] Oanh Nguyen and Van Vu. Random polynomials: central limit theorems for the real roots. *Duke Mathematical Journal*, 170(17):3745–3813, 2021.

[18] Oanh Nguyen and Van Vu. Roots of random functions: A framework for local universality. *American Journal of Mathematics*, 144(1):1–747, 2022.

[19] A. M. Odlyzko and B. Poonen. Zeros of polynomials with $0, 1$ coefficients. *Enseign. Math. (2)*, 39(3-4):317–348, 1993.

[20] Larry A Shepp and Robert J Vanderbei. The complex zeros of random polynomials. *Transactions of the American Mathematical Society*, pages 4365–4384, 1995.

Department of Mathematics, The Ohio State University, 231 W 18th Ave, Columbus, OH 43210 USA

Email address: nguyen.1261@osu.edu

Division of Applied Mathematics, Brown University, Providence, RI 02906, USA

Email address: oanh_nguyen1@brown.edu