ON RANDOM POLYNOMIALS WITH AN INTERMEDIATE NUMBER OF REAL ROOTS

MARCUS MICHELEN AND SEAN O'ROURKE

(Communicated by Zhen-Qing Chen)

ABSTRACT. For each $\alpha \in (0,1)$, we construct a bounded monotone deterministic sequence $(c_k)_{k \geqslant 0}$ of real numbers so that the number of real roots of the random polynomial $f_n(z) = \sum_{k=0}^{n} c_k \varepsilon_k z^k$ is $n^{\alpha + o(1)}$ with probability tending to one as the degree n tends to infinity, where (ε_k) is a sequence of i.i.d. (real) random variables of finite mean satisfying a mild anti-concentration assumption. In particular, this includes the case when (ε_k) is a sequence of i.i.d. standard Gaussian or Rademacher random variables. This result confirms a conjecture of O. Nguyen from 2019. More generally, our main results also describe several statistical properties for the number of real roots of f_n, including the asymptotic behavior of the variance and a central limit theorem.

1. INTRODUCTION

Consider the random polynomial

$$(1) \qquad f_n(z) = \sum_{k=0}^{n} c_k \varepsilon_k z^k,$$

where (c_k) is a deterministic sequence of real numbers and (ε_k) is a sequence of independent and identically distributed (i.i.d.) real-valued random variables (e.g., (ε_k) is a sequence of i.i.d. standard Gaussian random variables). Beginning with Bloch and Pólya's work in the 1930s [2], researchers have studied the roots (both real and complex) of f_n and used random polynomials for a variety of applications to many different areas of mathematics [1, 2, 4, 6, 16]. Among the earliest questions is how the number of real roots of f_n depends on the sequence (c_k).

In the case when (ε_k) is a sequence of i.i.d. Gaussian random variables, the Kac–Rice formula and other tools which take advantage of special properties of the Gaussian distribution are available and allow one to write the moments of the number of real roots as a deterministic integral; see, for instance, [4,8,9,14,15,24,25] and references therein. Recent developments have focused on universality results, where (ε_k) is a sequence of i.i.d. non-Gaussian random variables, and one often wishes to understand how the behavior of the roots depends of the distribution of (ε_k). We refer the reader to [3, 7, 11, 13, 17, 18, 20–22, 26, 27] and references therein; however, this list is far from complete.

©2024 American Mathematical Society

In this note, we focus on the number of real roots of f_n. Some of the most widely studied models for this problem include the Kac model ($c_k \equiv 1$) and the Weyl model ($c_k = 1/\sqrt{k!}$). Let R_n denote the number of real roots of f_n. We review some previously studied models here:

- In the case when (ε_k) is a sequence of i.i.d. standard Gaussian random variables and $c_k = 1$ for all $k \geqslant 1$, it was shown by Kac [14] that[1]

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

  where C is an explicit constant. See [5, 10, 12, 19] and the references therein for generalizations to other distributions.

- In the case when (ε_k) is a sequence of i.i.d. standard Gaussian random variables and $c_k = \frac{1}{\sqrt{k!}}$, it is known [4, 26] that

$$\mathbb{E}R_n = \left(\frac{2}{\pi} + o(1) \right) \sqrt{n}.$$

- When (ε_k) is a sequence of i.i.d. Rademacher random variables and $c_1 = 1/2$, $c_k = 1/(k!)^k$ for $k \geqslant 2$, Littlewood and Offord [16] showed that f_n has all real roots with probability 1.

- When $c_k = \exp(-k^\beta/2)$ and (ε_k) is a sequence of i.i.d. standard Gaussian random variables, Schehr and Majumdar [23] observed the following phase transition:

$$\mathbb{E}R_n = \begin{cases} \left(\frac{2}{\pi} + o(1)\right) \log n & \text{if } \beta \in [0,1) \\ \left(\frac{2}{\pi} + o(1)\right) \sqrt{\frac{\beta-1}{\beta}} n^{\beta/2} & \text{if } \beta \in (1,2) \\ (1 + o(1))n & \text{if } \beta > 2 \end{cases}.$$

  Notably, as β increases, the expected number of real roots "jumps" from order $\log n$ to $n^{1/2+\varepsilon}$ as β crosses 1.

Based on the above results, the following conjecture[2] is natural.

Conjecture 1.1 (O. Nguyen). *Let (ε_k) be an i.i.d. sequence of standard Gaussian random variables. Given $\alpha \in (0,1)$, there exists a bounded sequence (c_k) of real numbers so that $\mathbb{E}R_n = n^{\alpha + o(1)}$.*

Here, it is important that c_k only depends on k; if c_k is allowed to also depend on n, it becomes straightforward to establish Conjecture 1.1.

In this note, we confirm Conjecture 1.1 with a monotone decreasing sequence.

Theorem 1.2. *Let (ε_k) be an i.i.d. sequence of standard Gaussian random variables. For any $\alpha \in (0,1)$, there exists a monotone bounded sequence (c_k) of real numbers so that $\mathbb{E}R_n = n^{\alpha + o(1)}$.*

More generally, our main result below (Theorem 1.4) establishes several other statistical properties of R_n for our construction, including a central limit theorem, and shows these properties hold for many other coefficient distributions, beyond the Gaussian case.

---

[1]Here, $o(1)$ denotes a term which tends to zero as n tends to infinity; see Section 1.2 for more details about the asymptotic notation used here and throughout the paper.

[2]The conjecture was formulated in a workshop at the American Institute of Mathematics and can be found online at `http://aimpl.org/randpolyzero/`.

We now turn to the construction of the deterministic sequence (c_k). Fix $\alpha \in (0, 1)$. Define the sequence $m_j := 2\lfloor j^{1/\alpha} \rfloor$ and the interval of integers $I_j := (m_{j-1}, m_j]$; we set $m_{-1} = -1$ so that way the sequence $(I_j)_{j \geqslant 0}$ forms a partition of $\{0, 1, 2, \ldots\}$. We now define the sequence $(c_k)_{k \geqslant 0}$ via

$$(2) \qquad c_k = \exp(-2^j) \qquad \text{if} \qquad k \in I_j.$$

Note that the sequence (c_k) depends on α, but we do not denote this dependence in our notation.

We will assume (ε_k) is a sequence of i.i.d. random variables satisfying the following properties.

Assumption 1.3. Let $(\varepsilon_k)_{k \geqslant 0}$ be a sequence of real-valued i.i.d. random variables which satisfy the following assumptions:

    (i) $\mathbb{P}(\varepsilon_k > 0) =: p \in (0, 1)$
    (ii) $\mathbb{E}|\varepsilon_k| \leqslant 1$
    (iii) There exists a constant $C_0 > 0$ so that $\mathbb{P}(|\varepsilon_k| \leqslant t) \leqslant C_0 t$ for all $t \geqslant 0$.

Since the number of real roots is invariant under scaling the polynomial, Condition (ii) of Assumption 1.3 simply asserts that the random variables lie in L^1. Condition (iii) implies that ε_k is non-zero with probability 1. In particular, both the standard Gaussian distribution and the Rademacher distribution satisfy Assumption 1.3. We note that Condition (ii) and (iii) can be relaxed considerably, although we do not pursue this direction.

Fix $\alpha \in (0, 1)$. Define the random polynomial f_n of degree n by (1), where (c_k) is the sequence defined in (2) and (ε_j) is a sequence of i.i.d. real-valued random variables satisfying Assumption 1.3. Note that f_n has $(1 + o(1))(n/2)^\alpha$ "blocks" of coefficients, i.e., if we set $j_* = \max\{j : I_j \cap \{0, 1, \ldots, n\} \neq \varnothing\}$ then $j_* = (1 + o(1))(n/2)^\alpha$. The idea will be that the terms indexed by the sequence $(m_j)_{j \geqslant 0}$ will dictate the behavior of the polynomial with high probability. As such, we will let S_n denote the number of sign changes among the sequence $(\varepsilon_{m_j})_{j=0}^{j_*-1} \cup (\varepsilon_n)$ (note that by Condition (iii) ε_j is non-zero with probability 1).

We say an event $\mathcal{E}$ (which depends on n) holds with *overwhelming probability* if for every $\kappa > 0$ there exists a constant $C_\kappa > 0$ so that $\mathbb{P}(\mathcal{E}) \geqslant 1 - C_\kappa n^{-\kappa}$ for all $n > C_\kappa$.

Theorem 1.4. *Fix $\alpha \in (0, 1)$. Define the random polynomial f_n by (1), where (c_k) is the sequence defined in (2) and (ε_k) is a sequence of i.i.d. random variables satisfying Assumption 1.3. Let R_n be the number of real roots of f_n, and let S_n be the number of sign changes among the sequence $(\varepsilon_{m_j})_{j=0}^{j_*-1} \cup (\varepsilon_n)$. Then there exists an absolute constant $C > 0$ so that the event $|R_n - 2S_n| \leqslant C(\log n)^{1/\alpha}$ holds with overwhelming probability. In particular, there are constants c_p and c'_p depending only on p from Condition* (i) *so that $\mathbb{E}R_n \sim c_p(n/2)^\alpha$, $\operatorname{Var} R_n \sim c'_p(n/2)^\alpha$, $R_n/(c_p(n/2))^\alpha \to 1$ almost surely as $n \to \infty$, and*

$$\frac{R_n - \mathbb{E}R_n}{\sqrt{\operatorname{Var} R_n}} \longrightarrow \mathcal{N}(0, 1)$$

in distribution as $n \to \infty$, where $\mathcal{N}(0, 1)$ denotes the standard normal distribution.

Theorem 1.2 follows immediately from Theorem 1.4. We outline the proof in the following short subsection and prove Theorem 1.4 in Section 2.

1.1. **Outline and sketch of the proof of Theorem 1.4.** The intuition for why f_n has $n^{\alpha+o(1)}$ many real roots is that for each $t \in \mathbb{R}$ we have one of two types of behavior: either the m_j term dominates, in which case we do not have a real root with high probability; or we are transitioning from when the m_j term dominates to when the m_{j+1} term dominates, in which case we have a real root if and only if there is a sign change among $\{\varepsilon_{m_j}, \varepsilon_{m_{j+1}}\}$ with high probability. Since we have chosen m_j to always be an even integer, the behavior on the negative real axis is quite similar to that on the positive real axis, and so we only focus on the positive real axis.

In order to confirm this heuristic, we switch to exponential coordinates and define $g_n(t) = f_n(e^t)$ and first write a deterministic lemma that drives all of our comparisons between terms (Lemma 2.1). Here we introduce disjoint intervals $[a_j, b_j]$ where the m_j term of g_n dominates on $[a_j, b_j]$. We then convert our deterministic Lemma 2.1 to a probabilistic statement showing that in fact the m_j term of g_n does dominate on $[a_j, b_j]$ with high probability (Lemma 2.2). This immediately implies there are no real roots in $[a_j, b_j]$ with high probability (Lemma 2.3) and that if $\{\varepsilon_{m_j}, \varepsilon_{m_{j+1}}\}$ are the same sign then there is no real root in $[b_j, a_{j+1}]$ with high probability (Lemma 2.4). To handle the case when we *do* have a sign change among $\{\varepsilon_{m_j}, \varepsilon_{m_{j+1}}\}$, Lemma 2.2 shows that g_n has a sign change on $[b_j, a_{j+1}]$ with high probability thus showing that there is a real root; in order to show that there is only one real root, we show that the derivative of a rescaled version of g_n is nonzero on $[b_j, a_{j+1}]$ with high probability (Lemma 2.5). Finally, a simple application of Rouché's Theorem shows that there are not too many real roots near the origin (Lemma 2.6), allowing us to use the asymptotic results we have proven up until that point.

1.2. **Notation.** Starting in Section 2, unless otherwise noted, we use asymptotic notation under the assumption that $j \to \infty$; this differs from Section 1, where we used asymptotic notation under the assumption that $n \to \infty$. For two sequences (a_j) and (b_j), we write $a_j = O(b_j)$ if there exists a constant $C > 0$ so that $|a_j| \leqslant C b_j$ for all $j > C$. If the constant C depends on other parameters, e.g., $C = C_\alpha$, we denote this with subscripts, e.g., $a_j = O_\alpha(b_j)$. We write $a_j = o(b_j)$ to denote that $\lim_{j\to\infty} a_j/b_j = 0$. The notation $a_j \sim b_j$ denotes that $a_j = (1 + o(1))b_j$. We use i as an index (i.e., i does not denote the imaginary unit). We let C, c be positive constants (i.e., they do not depend on n or j) that may change from one occurrence to the next.

2. Proof of Theorem 1.4

It will often be more convenient to work in exponential coordinates, and we write $g_n(t) := f_n(e^t)$. Fix $\alpha \in (0, 1)$ and set $\beta := \min\{1/2, \frac{1-\alpha}{2\alpha}\}$. For each $j \geqslant 1$, define

$$a_j := \alpha j^{1-1/\alpha} 2^{j-2}(1 + j^{-\beta}) \qquad \text{and} \qquad b_j := \alpha j^{1-1/\alpha} 2^{j-1}(1 - j^{-\beta}).$$

It follows that that there exists a constant $C > 0$ (depending only on α) so that $a_j < b_j$ for all $j > C$ and $(a_j)_{j>C}$ and $(b_j)_{j>C}$ are increasing sequences. Further define the functions $\varphi_j(t) := c_{m_j} e^{t m_j}$ and set $j_* := \max\{j : I_j \cap \{0, 1, \ldots, n\} \neq \varnothing\}$ so we may write

$$g_n(t) = \sum_{j<j_*} \varphi_j(t) \sum_{i=0}^{m_j-m_{j-1}-1} \varepsilon_{m_j-i} e^{-it} + \sum_{i=m_{j_*-1}+1}^{n} c_i \varepsilon_i e^{it}.$$

We write g_n in this form to note that the blocks for $j < j_*$ are "complete" while the last block may potentially be "incomplete." We prove Theorem 1.4 via a sequence of lemmas, starting with a deterministic lemma showing that φ_j dominates on the interval $[a_j, b_j]$:

Lemma 2.1. *There are constants $C, c > 0$ depending only on α so that for all $j \geqslant 1$ we have*

$$(3) \qquad \log\left(\frac{\varphi_{j-1}(t)}{\varphi_j(t)}\right) \leqslant -c2^j j^{-\beta} + C$$

for any $t \geqslant a_j$ and

$$(4) \qquad \log\left(\frac{\varphi_{j+1}(t)}{\varphi_j(t)}\right) \leqslant -c2^j j^{-\beta} + C$$

for all $t \leqslant b_j$. In particular, there exist constants $C', c' > 0$ depending only on α so that

$$(5) \qquad \log\left(\frac{\varphi_{j-1}(t)}{\varphi_j(t)}\right) \leqslant -c'2^j j^{-\beta} \quad \text{and} \quad \log\left(\frac{\varphi_{j+1}(t)}{\varphi_j(t)}\right) \leqslant -c'2^j j^{-\beta}$$

for all $j \geqslant C'$, where the first bound holds for all $t \geqslant a_j$ and the second for all $t \leqslant b_j$.

Proof. For $t \geqslant a_j$, we compute

$$\log\left(\frac{\varphi_{j-1}(t)}{\varphi_j(t)}\right) \leqslant -2^{j-1} + 2^j + a_j\left(2\lfloor (j-1)^{1/\alpha}\rfloor - 2\lfloor j^{1/\alpha}\rfloor\right)$$

$$= 2^{j-1} - \left(\alpha j^{1-1/\alpha}2^{j-2}(1+j^{-\beta})\right)\left(\frac{2j^{1/\alpha-1}}{\alpha} + O_\alpha(1+j^{1/\alpha-2})\right)$$

$$= -2^{j-1}\left(j^{-\beta} + O_\alpha(j^{1-1/\alpha} + j^{-1})\right).$$

By the choice of β we note that $\beta < 1$ and $\beta < 1/\alpha - 1$ and so we see that there are constants $C, c > 0$ so that (3) holds. Similarly, for $t \leqslant b_j$, we compute

$$\log\left(\frac{\varphi_{j+1}(t)}{\varphi_j(t)}\right) \leqslant 2^j - 2^{j+1} + b_j\left(2\lfloor (j+1)^{1/\alpha}\rfloor - 2\lfloor j^{1/\alpha}\rfloor\right)$$

$$= -2^j + \left(\alpha j^{1-1/\alpha}2^{j-1}(1-j^{-\beta})\right)\left(\frac{2j^{1/\alpha-1}}{\alpha} + O_\alpha(1+j^{1/\alpha-2})\right)$$

$$= -2^j\left(j^{-\beta} + O_\alpha(j^{1-1/\alpha} + j^{-1})\right).$$

The choice of β again establishes (4). The bounds in (5) follow from (3) and (4). $\square$

We now extend this deterministic bound to show that the m_j term of our random polynomial dominates on $[a_j, b_j]$:

Lemma 2.2. *There exist constants $C, c > 0$ (depending only on α and C_0) so that the following holds.*

- *For any $j \geqslant 1$, with probability at least $1 - C\exp(-cj^2)$,*

$$(6) \qquad \min_{a_j \leqslant s \leqslant b_j}\left(\frac{c_{m_j}e^{m_j s}|\varepsilon_{m_j}|}{2} - \sum_{i \neq m_j}|\varepsilon_i||c_i|e^{is}\right) > 0.$$

- *For any $j \geqslant 1$, with probability at least $1 - C\exp(-cj^2)$,*

$$\inf_{a_j \leqslant s} \left(\frac{c_{m_j} e^{m_j s} |\varepsilon_{m_j}|}{2} - \sum_{i < m_j} |\varepsilon_i| c_i e^{is} \right) > 0. \tag{7}$$

- *For any $j \geqslant 1$, with probability at least $1 - C\exp(-cj^2)$,*

$$\inf_{s \leqslant b_j} \left(\frac{c_{m_j} e^{m_j s} |\varepsilon_{m_j}|}{2} - \sum_{i > m_j} |\varepsilon_i| c_i e^{is} \right) > 0. \tag{8}$$

- *With probability at least $1 - C\exp(-cj_*^2)$,*

$$\inf_{a_{j_*} \leqslant s} \left(c_n e^{ns} |\varepsilon_n| - \sum_{i < n} |\varepsilon_i| c_i e^{is} \right) > 0. \tag{9}$$

Proof. We prove (6); the proofs of (7), (8) and (9) are similar, and we omit the details. We first show that the m_j term dominates all others in its block. To this end, since $c_i = c_{m_j}$ for all $m_{j-1} + 1 \leqslant i \leqslant m_j$, we write

$$c_{m_j} e^{m_j s} |\varepsilon_{m_j}|/4 - \sum_{i=m_{j-1}+1}^{m_j-1} |\varepsilon_i| c_i e^{is} \overset{\mathrm{d}}{=} c_{m_j} e^{m_j s} \left(\frac{|\varepsilon_{m_j}|}{4} - \sum_{i=1}^{m_j - m_{j-1} - 1} |\varepsilon_i| e^{-is} \right),$$

where $\overset{\mathrm{d}}{=}$ denotes equality in distribution. Using Condition (ii), we bound

$$\mathbb{E} \max_{a_j \leqslant s \leqslant b_j} \sum_{i=1}^{m_j - m_{j-1} - 1} |\varepsilon_i| e^{-is} \leqslant 2e^{-a_j} = e^{-\Omega_\alpha(j^2)}.$$

By Markov's inequality along with Condition (iii), we can bound

$$\mathbb{P} \left(\min_{a_j \leqslant s \leqslant b_j} \left(c_{m_j} e^{m_j s} \frac{|\varepsilon_{m_j}|}{4} - \sum_{i \in I_j \setminus j} |\varepsilon_i| c_i e^{is} \right) \leqslant 0 \right) \leqslant \mathbb{P}(|\varepsilon_{m_j}| \leqslant C e^{-cj^2/2}) \tag{10}$$

$$+ e^{-\Omega_\alpha(j^2)}$$

$$= e^{-\Omega_\alpha(j^2)}.$$

To handle the coefficients not in I_j, first note that

$$\sum_{i \notin I_j} |\varepsilon_i| \frac{c_i e^{is}}{c_{m_j} e^{m_j s}} \leqslant \sum_{r \neq j} \frac{\varphi_r(s)}{\varphi_j(s)} \sum_{i \in I_r} |\varepsilon_i|.$$

By iteratively applying Lemma 2.1 and using Condition (ii), we have

$$\mathbb{E} \max_{a_j \leqslant s \leqslant b_j} \left(\sum_{r \neq j} \frac{\varphi_r(s)}{\varphi_j(s)} \sum_{i \in I_r} |\varepsilon_i| \right) \leqslant C m_j e^{-cj^2} + \sum_{r > j} C |I_r| e^{-cr^2}$$

$$= e^{-\Omega_\alpha(j^2)}.$$

By applying Markov's inequality and Condition (iii) as in the proof of (10), one finds

$$\mathbb{P} \left(\min_{a_j \leqslant s \leqslant b_j} \left(\frac{c_{m_j} e^{m_j s} |\varepsilon_{m_j}|}{4} - \sum_{i \notin I_j} |\varepsilon_i| c_i e^{is} \right) \leqslant 0 \right) = e^{-\Omega_\alpha(j^2)}. \tag{11}$$

Combining bounds (10) and (11) shows (6). $\qquad\square$

We now want to show that the number of positive real roots is the same as the number of sign changes among the coefficients $(\varepsilon_{m_j})_j$. We first show that there are no real roots in the interval $[a_j, b_j]$ with high probability.

Lemma 2.3. *There exist constants $C, c > 0$ (depending only on α and C_0) so that the following holds. For each $j \geqslant 1$ with $m_j \leqslant n$, with probability at least $1 - Ce^{-cj^2}$ we have that $\mathrm{sgn}(g_n(s)) = \mathrm{sgn}(\varepsilon_{m_j})$ for all $s \in [a_j, b_j]$. In particular, with probability at least $1 - Ce^{-cj^2}$, $g_n(s)$ has no roots in $[a_j, b_j]$. Similarly, for all $s \geqslant a_{j_*}$ we have $\mathrm{sgn}(g_n(s)) = \mathrm{sgn}(\varepsilon_n)$ with probability at least $1 - Ce^{-cj_*^2}$.*

Proof. Applying the triangle inequality, we see that

$$|g_n(s)| \geqslant c_{m_j} e^{m_j s} |\varepsilon_{m_j}| - \sum_{i \neq m_j} |\varepsilon_i| c_i e^{is}.$$

Thus, the conclusion follows from Lemma 2.2. The assertion for $s \geqslant a_{j_*}$ follows from Lemma 2.2 as well. $\qquad\square$

We next show that if there is no sign change among $\{\varepsilon_{m_j}, \varepsilon_{m_{j+1}}\}$ then there is no real root on $[b_j, a_{j+1}]$.

Lemma 2.4. *There exist constants $C, c > 0$ (depending only on α and C_0) so that the following holds. For each $j < j_*$ with $m_{j+1} \leqslant n$, with probability at least $1 - Ce^{-cj^2}$, if $\mathrm{sgn}(\varepsilon_{m_j}) = \mathrm{sgn}(\varepsilon_{m_{j+1}})$ then $g_n(t)$ has no roots in the interval $[b_j, a_{j+1}]$. Similarly, if $m_{j_*} > n$ and $\mathrm{sgn}(\varepsilon_{m_{j_*-1}}) = \mathrm{sgn}(\varepsilon_n)$ then g_n has no roots in the interval $[b_{j_*-1}, a_{j_*}]$ with probability at least $1 - Ce^{-cj_*^2}$.*

Proof. Assume without loss of generality that $\mathrm{sgn}(\varepsilon_{m_j}) = \mathrm{sgn}(\varepsilon_{m_{j+1}}) = +1$. Then for any $s \in [b_j, a_{j+1}]$, we bound

$$g_n(s) \geqslant \left(\frac{\varepsilon_{m_j} c_{m_j} e^{m_j s}}{2} - \sum_{i < m_j} |\varepsilon_i| c_i e^{is} \right) + \left(\frac{\varepsilon_{m_{j+1}} c_{m_{j+1}} e^{m_{j+1} s}}{2} - \sum_{i > m_{j+1}} |\varepsilon_i| c_i e^{is} \right)$$

$$+ \left(\frac{\varepsilon_{m_j} c_{m_j} e^{m_j s}}{2} + \frac{\varepsilon_{m_{j+1}} c_{m_{j+1}} e^{m_{j+1} s}}{2} - \sum_{i = m_j + 1}^{m_{j+1} - 1} |\varepsilon_i| c_i e^{is} \right)$$

The first two terms can be bounded using (7) and (8) from Lemma 2.2. The third term on the right-hand side can be bounded by following a nearly identical argument which led to (10). The assertion for $m_{j_*} > n$ follows in a similar way from Lemma 2.2. $\qquad\square$

We now handle the other case and show that if there is a sign change among $\{\varepsilon_{m_j}, \varepsilon_{m_{j+1}}\}$ then there is precisely one real root on $[b_j, a_{j+1}]$. For this, we show that after rescaling our function g_n, the derivative does not have a sign change with high probability.

Lemma 2.5. *There exist constants $C, c > 0$ (depending only on α and C_0) so that the following holds. For each $j \geqslant 1$ with $m_{j+1} \leqslant n$, with probability at least $1 - Ce^{-cj^2}$, if $\mathrm{sgn}(\varepsilon_{m_j}) = -\mathrm{sgn}(\varepsilon_{m_{j+1}})$, then g_n has a single root in the interval $[b_j, a_{j+1}]$. Similarly, if $m_{j_*} > n$ and $\mathrm{sgn}(\varepsilon_{m_{j_*-1}}) = -\mathrm{sgn}(\varepsilon_n)$ then g_n has a single root in the interval $[b_{j_*-1}, a_{j_*}]$ with probability at least $1 - Ce^{-cj_*^2}$.*

Proof. We prove the first assertion as the second is similar. By replacing g_n with $-g_n$ if necessary, assume without loss of generality that $\varepsilon_{m_{j+1}} > 0$ and $\varepsilon_{m_j} < 0$. Write $g_n(t) = e^{t(m_{j+1}+m_j)/2} h_n(t)$ with

$$h_n(t) := \varepsilon_{m_j} c_{m_j} e^{-t(m_{j+1}-m_j)/2} + \varepsilon_{m_{j+1}} c_{m_{j+1}} e^{t(m_{j+1}-m_j)/2}$$
$$+ \sum_{i \notin \{m_j, m_{j+1}\}} \varepsilon_i c_i e^{t(i - m_{j+1}/2 - m_j/2)} \, .$$

It is sufficient to show that $h_n(t)$ has precisely one root in $[b_j, a_{j+1}]$. By Lemma 2.3, h_n exhibits a sign change on the interval $[b_j, a_{j+1}]$ with probability at least $1 - Ce^{-cj^2}$. Thus, it will be sufficient to show that with probability at least $1 - Ce^{-cj^2}$, h'_n is positive on the interval. We compute

$$\frac{2e^{t(m_{j+1}+m_j)/2}}{m_{j+1} - m_j} h'_n(t) = (-\varepsilon_{m_j}) c_{m_j} e^{m_j t} + \varepsilon_{m_{j+1}} c_{m_{j+1}} e^{m_{j+1} t}$$
$$+ \sum_{i \notin \{m_j, m_{j+1}\}} \varepsilon_i c_i e^{it} \cdot \frac{2i - m_{j+1} - m_j}{m_{j+1} - m_j} \, .$$

Recalling that $(-\varepsilon_{m_j}) > 0$ and $\varepsilon_{m_{j+1}} > 0$, it suffices to show that

$$|\varepsilon_{m_j}| c_{m_j} e^{m_j t} - \sum_{i < m_j} |\varepsilon_i| c_i e^{it} \left| \frac{2i - m_{j+1} - m_j}{m_{j+1} - m_j} \right| > 0$$

and

$$|\varepsilon_{m_{j+1}}| c_{m_{j+1}} e^{m_{j+1} t} - \sum_{i > m_j, i \neq m_{j+1}} |\varepsilon_i| c_i e^{it} \left| \frac{2i - m_{j+1} - m_j}{m_{j+1} - m_j} \right| > 0$$

with probability at least $1 - Ce^{-cj^2}$. Since

$$\left| \frac{2i - m_{j+1} - m_j}{m_{j+1} - m_j} \right| \leqslant 2i + 2m_j$$

(this follows from the triangle inequality since $m_{j+1} - m_j \geqslant 1$), the conclusion follows by applying the proof of Lemma 2.2 (with only slight changes to account for this extra factor of $2i + 2m_j$). $\qquad\square$

Since our previous lemmas are all asymptotic as $j \to \infty$, we now need to show that the behavior of the first few terms does not contribute too many roots close to the origin. A simple application of Rouché's theorem will allow us to bound the number of roots in a ball centered at 0:

Lemma 2.6. *There exist constants $C, c > 0$ (depending only on α and C_0) so that the following holds. For each $j \geqslant 1$ with $m_j < n$, with probability at least $1 - Ce^{-cj^2}$, the polynomial $f_n(z)$ has at most m_j roots in the disk $\{z \in \mathbb{C} : |z| \leqslant e^{b_j}\}$.*

Proof. We will use Rouché's theorem. In particular, if we show that for all $z \in \mathbb{C}$ with $|z| = e^{b_j}$ we have

$$(12) \qquad\qquad |\varepsilon_{m_j} c_{m_j} z^{m_j}| > \left| \sum_{i \neq m_j} \varepsilon_i c_i z^i \right| ,$$

then we will have that $f_n(z)$ has at most m_j roots in the disk $\{z \in \mathbb{C} : |z| \leqslant e^{b_j}\}$. We lower bound

$$\min_{|z|=e^{b_j}} \left(|\varepsilon_{m_j} c_{m_j} z^{m_j}| - \left| \sum_{i \neq m_j} \varepsilon_i c_i z^i \right| \right) \geqslant |\varepsilon_{m_j}| c_{m_j} e^{m_j b_j} - \sum_{i \neq m_j} |\varepsilon_i| c_i e^{i b_j}$$

and hence the conclusion follows from Lemma 2.2. $\qquad\square$

With the lemmas above in hand, we are now ready to complete the proof of Theorem 1.4.

Proof of Theorem 1.4. Choose $k = \log n$ and apply Lemma 2.6 to see that, with overwhelming probability, the number of real roots of f_n differs from the number of real roots in $\{z \in \mathbb{R} : |z| > e^{b_k}\}$ by at most $2(\log n)^{1/\alpha}$. Combining Lemmas 2.3, 2.4 and 2.5 shows that, with overwhelming probability, the number of real roots in $\{t \in \mathbb{R} : t > e^{b_k}\}$ is equal to the number of sign changes of the sequence $(\varepsilon_{m_j})_{j=k}^{j_*-1} \cup (\varepsilon_n)$; similarly, the same argument shows that the number of *negative* real roots in $\{t \in \mathbb{R} : t < -e^{b_k}\}$ is also equal to the number of sign changes of $(\varepsilon_{m_j})_{j=k}^{j_*-1} \cup (\varepsilon_n)$.

This shows that $|R_n - 2S_n| \leqslant 4(\log n)^{1/\alpha}$ with overwhelming probability. The other conclusions of Theorem 1.4 follow from this bound. For example, the classical law of large numbers for S_n implies the same conclusion for R_n. $\qquad\square$

ACKNOWLEDGEMENT

REFERENCES

[1] A. T. Bharucha-Reid and M. Sambandham, *Random polynomials*, Probability and Mathematical Statistics, Academic Press, Inc., Orlando, FL, 1986. MR856019 ↑4933

[2] A. Bloch and G. Pólya, *On the Roots of Certain Algebraic Equations*, Proc. London Math. Soc. (2) **33** (1931), no. 2, 102–114, DOI 10.1112/plms/s2-33.1.102. MR1576817 ↑4933

[3] Yen Do, Oanh Nguyen, and Van Vu, *Roots of random polynomials with coefficients of polynomial growth*, Ann. Probab. **46** (2018), no. 5, 2407–2494, DOI 10.1214/17-AOP1219. MR3846831 ↑4933

[4] Alan Edelman and Eric Kostlan, *How many zeros of a random polynomial are real?*, Bull. Amer. Math. Soc. (N.S.) **32** (1995), no. 1, 1–37, DOI 10.1090/S0273-0979-1995-00571-9. MR1290398 ↑4933, 4934

[5] Paul Erdös and A. C. Offord, *On the number of real roots of a random algebraic equation*, Proc. London Math. Soc. (3) **6** (1956), 139–160, DOI 10.1112/plms/s3-6.1.139. MR73870 ↑4934

[6] P. J. Forrester and G. Honner, *Exact statistical properties of the zeros of complex random polynomials*, J. Phys. A **32** (1999), no. 16, 2961–2981, DOI 10.1088/0305-4470/32/16/006. MR1690355 ↑4933

[7] 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), no. Veroyatnost′i Statistika. 24, 102–111, DOI 10.1007/s10958-018-3705-4; English transl., J. Math. Sci. (N.Y.) **229** (2018), no. 6, 664–670. MR3602403 ↑4933

[8] J. M. Hammersley, *The zeros of a random polynomial*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II, Univ. California Press, Berkeley-Los Angeles, Calif., 1956, pp. 89–111. MR84888 ↑4933

[9] J. Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág, *Zeros of Gaussian analytic functions and determinantal point processes*, University Lecture Series, vol. 51, American Mathematical Society, Providence, RI, 2009, DOI 10.1090/ulect/051. MR2552864 ↑4933

[10] I. A. Ibragimov and N. B. Maslova, *The average number of real roots of random polynomials* (Russian), Dokl. Akad. Nauk SSSR **199** (1971), 13–16. MR292134 ↑4934

[11] I. A. Ibragimov and N. B. Maslova, *The mean number of real zeros of random polynomials. I. Coefficients with zero mean* (Russian, with English summary), Teor. Verojatnost. i Primenen. **16** (1971), 229–248. MR286157 ↑4933

[12] I. A. Ibragimov and N. B. Maslova, *The mean number of real zeros of random polynomials. I. Coefficients with zero mean* (Russian, with English summary), Teor. Verojatnost. i Primenen. **16** (1971), 229–248. MR286157 ↑4934

[13] Zakhar Kabluchko and Dmitry Zaporozhets, *Asymptotic distribution of complex zeros of random analytic functions*, Ann. Probab. **42** (2014), no. 4, 1374–1395, DOI 10.1214/13-AOP847. MR3262481 ↑4933

[14] M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc. **49** (1943), 314–320, DOI 10.1090/S0002-9904-1943-07912-8. MR7812 ↑4933, 4934

[15] M. Kac, *On the Average Number of Real Roots of a Random Algebraic Equation (II)*, Proc. London Math. Soc. (2) **50** (1949), no. 6, 390–408, DOI 10.1112/plms/s2-50.6.401. MR1575245 ↑4933

[16] J. E. Littlewood and A. C. Offord, *On the Number of Real Roots of a Random Algebraic Equation*, J. London Math. Soc. **13** (1938), no. 4, 288–295, DOI 10.1112/jlms/s1-13.4.288. MR1574980 ↑4933, 4934

[17] B. F. Logan and L. A. Shepp, *Real zeros of random polynomials*, Proc. London Math. Soc. (3) **18** (1968), 29–35, DOI 10.1112/plms/s3-18.1.29. MR234512 ↑4933

[18] B. F. Logan and L. A. Shepp, *Real zeros of random polynomials. II*, Proc. London Math. Soc. (3) **18** (1968), 308–314, DOI 10.1112/plms/s3-18.2.308. MR234513 ↑4933

[19] Hoi Nguyen, Oanh Nguyen, and Van Vu, *On the number of real roots of random polynomials*, Commun. Contemp. Math. **18** (2016), no. 4, 1550052, 17, DOI 10.1142/S0219199715500522. MR3493213 ↑4934

[20] Oanh Nguyen and Van Vu, *Roots of random functions: a framework for local universality*, Amer. J. Math. **144** (2022), no. 1, 1–74, DOI 10.1353/ajm.2022.0000. MR4367414 ↑4933

[21] Igor E. Pritsker, *Zero distribution of random polynomials*, J. Anal. Math. **134** (2018), no. 2, 719–745, DOI 10.1007/s11854-018-0023-1. MR3771497 ↑4933

[22] Igor E. Pritsker and Aaron M. Yeager, *Zeros of polynomials with random coefficients*, J. Approx. Theory **189** (2015), 88–100, DOI 10.1016/j.jat.2014.09.003. MR3280673 ↑4933

[23] Grégory Schehr and Satya N. Majumdar, *Condensation of the roots of real random polynomials on the real axis*, J. Stat. Phys. **135** (2009), no. 4, 587–598, DOI 10.1007/s10955-009-9755-8. MR2544105 ↑4934

[24] Larry A. Shepp and Robert J. Vanderbei, *The complex zeros of random polynomials*, Trans. Amer. Math. Soc. **347** (1995), no. 11, 4365–4384, DOI 10.2307/2155041. MR1308023 ↑4933

[25] Mikhail Sodin, *Zeroes of Gaussian analytic functions*, European Congress of Mathematics, Eur. Math. Soc., Zürich, 2005, pp. 445–458. MR2185759 ↑4933

[26] Terence Tao and Van Vu, *Local universality of zeroes of random polynomials*, Int. Math. Res. Not. IMRN **13** (2015), 5053–5139, DOI 10.1093/imrn/rnu084. MR3439098 ↑4933, 4934

[27] D. I. Šparo and M. G. Šur, *On the distribution of roots of random polynomials* (Russian, with English summary), Vestnik Moskov. Univ. Ser. I Mat. Meh. **1962** (1962), no. 3, 40–43. MR139199 ↑4933

Department of Mathematics, Statistics and Computer Science, University of Illinois, Chicago, 851 S. Morgan St., Chicago, Illinois 60607
    Email address: michelen@uic.edu

Department of Mathematics, University of Colorado Boulder, Campus Box 395, Boulder, Colorado 80309-0395
    Email address: sean.d.orourke@colorado.edu