PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 00, Number 0, Pages 000–000
S 0002-9939(XX)0000-0

EXPECTED DISCREPANCY FOR ZEROS OF RANDOM ALGEBRAIC POLYNOMIALS

IGOR E. PRITSKER AND ALAN A. SOLA

Dedicated to Vladimir Andrievskii on his 60th birthday

Abstract. We study asymptotic clustering of zeros of random polynomials, and show that the expected discrepancy of roots of a polynomial of degree n, with not necessarily independent coefficients, decays like $\sqrt{\log n/n}$. Our proofs rely on discrepancy results for deterministic polynomials, and order statistics of a random variable. We also consider the expected number of zeros lying in certain subsets of the plane, such as circles centered on the unit circumference, and polygons inscribed in the unit circumference.

1. Introduction

1.1. **Random polynomials and their zeros.** Let $\{C_k\}_{k=0}^\infty$ be a sequence of independent and identically distributed (iid) complex-valued random variables. In this paper, we study families of random polynomials

$$P_n(z) = \sum_{k=0}^{n} C_k z^k \tag{1.1}$$

and the geometry of their zeros; we let $\mathcal{Z}(P_n) = \{Z_1, \ldots, Z_n\}$ denote the set of complex zeros of such a polynomial of degree n. We use the notation $P_n(z) = \sum_{k=0}^{n} c_k z^k$ whenever we make statements that apply to any polynomial with $c_0, \ldots, c_n \in \mathbb{C}$ and reserve capital letters for random coefficients. The zeros $\{Z_k\}_{k=1}^n$ of a random polynomial P_n define a natural random measure on $\mathbb{C}$, the *counting measure of roots* or *empirical measure*

$$\tau_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{Z_k}.$$

Random polynomials of the form (1.1) have been studied by many authors, and it is known that, under mild conditions on the distribution of the

Received by the editors 8/18/2012.
2010 *Mathematics Subject Classification.* Primary: 30C15; Secondary: 30B20, 60B10.
Key words and phrases. Random polynomials, expected discrepancy of roots, expected number of roots in special sets.
Pritsker acknowledges support from NSA under grant H98230-12-1-0227; Sola acknowledges support from the EPSRC under grant EP/103372X/1.

coefficients of P_n, these empirical measures converge to $\mu_{\mathbb{T}}$, the normalized arc-length measure, in the weak* topology (or weakly, in the language of probability theory) as $n \to \infty$. For the history of the subject and a list of references we refer the reader to the books [5, 12]; we shall discuss some recent results shortly.

It is natural to ask how fast the empirical measures converge, and in this work, we provide estimates on the expected rate of convergence of associated quantities that measure the distance between counting measures and the uniform measure on the unit circle $\mathbb{T} = \{e^{i\theta} \colon \theta \in [0, 2\pi)\}$.

1.2. **Discrepancy of zeros and norms on polynomials.** For a polynomial P_n, we write $N(\alpha, \beta)$ for the number of elements of $\mathcal{Z}(P_n)$ that are contained in the sector

$$S(\alpha, \beta) = \{z \in \mathbb{C} : 0 < \alpha \le \arg z < \beta \le 2\pi\};$$

later on we shall also work with annular sectors of the form

$$A_r(\alpha, \beta) = \{z \in \mathbb{C} : r < |z| < 1/r,\ \alpha \le \arg z < \beta\}, \quad 0 < r < 1. \tag{1.2}$$

A very classical result concerning the *angular discrepancy of zeros* is the theorem of Erdős and Turán [11], which in its improved form due to Ganelius (see [14]) asserts that

$$\left|\frac{N(\alpha, \beta)}{n} - \frac{\beta - \alpha}{2\pi}\right| \le \sqrt{\frac{2\pi}{\mathbf{k}}}\sqrt{\frac{1}{n}\log\left[\frac{\|P_n\|_\infty}{\sqrt{|c_0 c_n|}}\right]}; \tag{1.3}$$

here $\mathbf{k} = \sum_{k=0}^{\infty}(-1)^k/(2k+1)^2$ denotes Catalan's constant, and $\|P_n\|_\infty = \sup_{\mathbb{T}}|P_n|$. Suppose now that the coefficients of P_n (random or deterministic) are uniformly bounded with $|c_k| \le K$, say. We then have the easy estimate

$$\log\|P_n\|_\infty \le \log\left(\max_{z \in \mathbb{T}} \sum_{k=0}^{n} |c_k z^k|\right) \le \log(n+1) + \log K,$$

and by (1.3), it follows that

$$\left|\frac{N(\alpha, \beta)}{n} - \frac{\beta - \alpha}{2\pi}\right| \le C\sqrt{\frac{\log(n+1) + \log K - \log\sqrt{|c_0 c_n|}}{n}}.$$

Thus we see that if $c_0 c_n \ne 0$ (almost surely), the discrepancy is of the order $\sqrt{\log n/n}$. Moreover, it can be shown that this rate of decay is essentially best possible in the sense that one can construct (deterministic) families of polynomials that exhibit discrepancy of this order (see [11, 2, 21]). For general coefficients however, it is hard to obtain effective estimates on $\|P_n\|_\infty$.

In order to state our results, we need to introduce certain additional norms on the circle. For a polynomial P_n and $0 < p < \infty$, we set

$$\|P_n\|_p = \left(\frac{1}{2\pi}\int_{\mathbb{T}} |P_n(e^{i\theta})|^p \, d\theta\right)^{1/p}.$$

When $P_n(z) = \sum_{k=0}^n c_k z^k$ and $p = 2$, we have $\|P_n\|_2^2 = \sum_{k=0}^n |c_k|^2$.

As usual, we define the Mahler measure (geometric mean) of P_n by

$$M(P_n) = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log|P_n(e^{i\theta})|d\theta\right),$$

and we set $m(P_n) = \log(M(P_n))$. We also use

$$M^+(P_n) = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log^+|P_n(e^{i\theta})|d\theta\right)$$

and $m^+(P_n) = \log(M^+(P_n))$. It is well known that

$$M(P_n) \le \|P_n\|_p \le \|P_n\|_q \le \|P_n\|_\infty, \quad 0 < p < q < \infty. \tag{1.4}$$

Furthermore, the definitions immediately give that

$$M(P_n) \le M^+(P_n) \le \|P_n\|_\infty. \tag{1.5}$$

Recent papers on random polynomials and the behavior of their roots include [22, 15, 16, 17, 18]. Improving an earlier result of Šparo and Šur, Ibragimov and Zaporozhets [17] prove that the condition $\mathbb{E}[\log^+|C_0|] < \infty$ is both necessary and sufficient for almost sure asymptotic concentration of roots on the unit circumference. The paper [18] deals with the interesting case of heavy-tailed coefficients, where $\mathbb{E}[\log^+|C_0|] = \infty$ and the asymptotic distribution of roots is uniform in argument, but the radial positions of the roots accumulate on more than one circle. In [15], Hughes and Nikeghbali deal with polynomials with not necessarily independent coefficients, and use estimates on $\mathbb{E}[\log(\sum_k |C_k|)]$ to deduce, via Erdős and Turán's theorem, that roots concentrate on the unit circumference.

1.3. **Overview of the paper.** In this paper, we initiate a systematic quantitative study of convergence results for general random polynomials with coefficients that are not necessarily bounded, but satisfy the condition for asymptotic concentration. Using rather elementary methods, similar in spirit to those in [15], we obtain results that are optimal unless further restrictions are introduced. If it is not stated otherwise, we shall impose the following standing assumption on the coefficients:

- $C_0, C_1, \ldots$ are independent and identically distributed (iid) complex random variables, with absolutely continuous distribution and $\mathbb{E}[|C_0|^t] = \mu < \infty$ for some $t > 0$.

We sometimes relax this assumption at the price of other more restrictive hypotheses.

Our main object of study are *expected discrepancies*, and we proceed as follows. We seek to control the discrepancy of zeros of a given random polynomial in annular sectors, first in terms of $m^+(P_n)$ and $m(P_n)$, and then $\log(\|P_n\|_2)$. To achieve this, we need to extend certain results of Mignotte and others; this is done in Section 2. When estimating the expected discrepancy of zeros, we find it convenient to consider the quantity $\mathbb{E}[\log(\max_k |C_k|)]$, since it appears in both upper and lower estimates on

$\mathbb{E}[\log(\|P_n\|_2)]$ and is amenable to elementary estimates. Under our standing assumption, we find that $\mathbb{E}[\log(\max_k |C_k|)] = O(\log n)$. More precise statements are given in Section 3. Our main result on the expected discrepancy is contained in Theorem 3.3, and it deals with radial and angular parts of zeros simultaneously:

$$\mathbb{E}\left[\left|\tau_n\left(A_r(\alpha,\beta)\right) - \frac{\beta-\alpha}{2\pi}\right|\right] \le C(r,t)\sqrt{\frac{\log(n+1)+\log\mu}{n}}$$

as $n \to \infty$; here, $C(r,t)$ is a constant that only depends on $r < 1$ and on $t > 0$. Inspired by the work of Borwein, Erdélyi, and Littmann [7], we also derive a number of corollaries concerning the *expected number of zeros* of random polynomials in polygons inscribed in the unit disk, and other natural sets. Under the additional assumption of a finite second moment, we indicate how our results extend to random coefficients that are not necessarily independent or identically distributed, yielding the same rate of decay for the expected discrepancy. Finally, in Section 4, we present some elementary examples that illustrate that while $\mathbb{E}[\log(\max_k |C_k|)]$ can be smaller in special cases, $O(\log n)$ is the correct order of magnitude if no additional assumptions are made.

Full proofs are given in Section 5.

2. Annular discrepancies and norms on $\mathbb{T}$

The theorem of Erdős and Turán has been extended by several authors (see, for instance [3, 2, 19, 20, 21]). In particular, Mignotte showed that one can use the weaker norm M^+ in estimates like (1.3), see [19] and [2]. It is also desirable to include information about the radial behavior of the roots. This can be easily done by using Jensen's formula, which leads to the following discrepancy estimate for the annular sectors (1.2).

Proposition 2.1. *Let* $P_n(z) = \sum_{k=0}^n c_k z^k$, $c_k \in \mathbb{C}$, *and assume* $c_0 c_n \neq 0$. *For any* $r \in (0,1)$ *and* $0 \le \alpha < \beta < 2\pi$, *we have*

$$\left|\tau_n\left(A_r(\alpha,\beta)\right) - \frac{\beta-\alpha}{2\pi}\right| \le \sqrt{\frac{2\pi}{\mathbf{k}}}\sqrt{\frac{1}{n}\, m^+\left(\frac{P_n}{\sqrt{|c_0 c_n|}}\right) + \frac{2}{n(1-r)}\, m\left(\frac{P_n}{\sqrt{|c_0 c_n|}}\right)}. \tag{2.1}$$

This estimate shows how close the zero counting measure τ_n is to $\mu_{\mathbb{T}}$. It is often more convenient to use the standard L^p norms, especially the L^2 norm. We mention the following elementary but useful estimate.

Proposition 2.2. *If* $p \in (0,\infty)$ *and* $\|P_n\|_p \ge 1$, *then*

$$m^+(P_n) \le \log\|P_n\|_p + 1/(ep).$$

3. Expected discrepancies for random polynomials

3.1. **Expectation of norms and the maximum of coefficients.** In order to estimate the *expected discrepancy* of zeros for classes of random polynomials, we apply Propositions 2.1 and 2.2, which requires an estimate on

$$\mathbb{E}[\log \|P_n\|_2] = \frac{1}{2}\mathbb{E}\left[\log\left(\sum_{k=0}^{n} |C_k|^2\right)\right].$$

Expected values of L^p norms and Mahler measures for random polynomials have been considered by a number of authors. For instance, Fielding (see [13]) has computed $\mathbb{E}[\log M(P_n)]$ for polynomials with coefficents uniformly distributed on $\mathbb{T}$; he obtains

$$\mathbb{E}[\log M(P_n)] = \frac{1}{2}\log(n+1) - \frac{\gamma}{2} + O(n^{-1/2+\delta})$$

for arbitrary $\delta > 0$. Here, γ denotes Euler's constant. Expected values of L^p norms of random polynomials have also been studied by Borwein and Lockhart, and Choi and Mossinghoff (see [8, 9]), among others.

It is elementary that

$$\max_{k=0,\dots,n} |C_k| \le \left(\sum_{k=0}^{n} |C_k|^p\right)^{1/p} \le (n+1)^{1/p} \max_{k=0,\dots,n} |C_k|,$$

and from this it follows that

$$\mathbb{E}[\log(\max_k |C_k|)] \le \mathbb{E}[\log \|P_n\|_2] \le \frac{1}{2}\log(n+1) + \mathbb{E}[\log(\max_k |C_k|)]. \tag{3.1}$$

Hence, we obtain an upper estimate for the discrepancy in terms of

$$\frac{1}{2}\log(n+1) + \mathbb{E}\left[\log \max_{k=0,\dots,n} |C_k|\right] - \mathbb{E}[\log |C_0|],$$

and thus the same order of decay as for bounded coefficients if we can show, for instance, that $\mathbb{E}\left[\log(\max_{k=0,\dots,n} |C_k|)\right] = O(\log n)$. On the other hand, the lower bound in (3.1) means that approaching the expected discrepancy via ℓ^q norms of coefficients (which bound $\|P_n\|_p$ from below for $1 \le p \le 2$) will not work if the expected value is too large; in view of Ibragimov and Zaporozhets' result [17], the roots do not exhibit (almost sure) asymptotic clustering if $\mathbb{E}[\log^+ |C_0|] = \infty$.

3.2. **Expectation of** $\log(\max_k |C_k|)$**.** We set $R_C(r) = \mathbb{P}(|C_0| \le r)$ and let $\rho_C(r) = R_C'(r)$, $r \in [0,\infty)$, denote the density of the non-negative random variable $|C_0|$, which exists and satisfies $\int_0^\infty r^t \rho_C(r)dr < \infty$ by our standing assumptions. Indeed, if $\rho(r,\theta)$ is the density of C_0 with respect to the area measure, then we can write $\mathbb{P}(|C_0| \le r) = \int_0^r \left(\int_0^{2\pi} s\,\rho(s,\theta)\,d\theta\right) ds = \int_0^r \rho_C(s)\,ds$. As is standard in order statistics (see [10]), we now express the density of the random variable $Y_n = \max_{k=0,\dots,n} |C_k|$ in terms of ρ_C.

Lemma 3.1. *Suppose $n \geq 1$. Then the density of $Y_n = \max_{k=0,\dots,n} |C_k|$ is given by*

$$\rho_{Y_n}(r) = (n+1)\rho_C(r)[R_C(r)]^n. \tag{3.2}$$

Using Lemma 3.1, we proceed by estimating

$$\mathbb{E}\left[\log Y_n\right] = \int_0^\infty (n+1)\rho_C(r)[R_C(r)]^n \log r \, dr,$$

or equivalently, with $x(u) = R_C^{-1}(u)$,

$$\mathbb{E}\left[\log Y_n\right] = \int_0^1 (n+1)\log(x(u))u^n \, du.$$

We recall that $\mathbb{E}[|C_0|^t] = \mu < \infty$, and Jensen's inequality now yields

Lemma 3.2. *We have*

$$\mathbb{E}\left[\log Y_n\right] \leq \frac{1}{t}(\log(n+1) + \log \mu). \tag{3.3}$$

3.3. **Main Results.** Combining Propositions 2.1 and 2.2 with (3.1), and using Lemma 3.2, we obtain the desired expected discrepancy result.

Theorem 3.3. *If the coefficients of $P_n(z) = \sum_{k=0}^n C_k z^k$ are iid complex random variables with absolutely continuous distribution and $\mathbb{E}[|C_0|^t] < \infty$, then we have for all large $n \in \mathbb{N}$ that*

$$\begin{aligned} &\mathbb{E}\left[\left|\tau_n\left(A_r(\alpha,\beta)\right) - \frac{\beta-\alpha}{2\pi}\right|\right] \\ &\leq \left(\sqrt{\frac{2\pi}{\mathbf{k}}} + \frac{2}{1-r}\right)\sqrt{\frac{\frac{t+2}{2t}\log(n+1) + \frac{1}{t}\log\mathbb{E}[|C_0|^t] + \frac{1}{2e} - \mathbb{E}[\log|C_0|]}{n}}. \end{aligned} \tag{3.4}$$

Note that for any set $E \subset \mathbb{C}$, the number of zeros of the polynomial P_n in E is given by $n\tau_n(E)$. Thus our results may be stated in terms of the expected number of zeros in certain sets (see [1, Ch. 5] for a general theorem). We give several examples of such statements below. We point out that for some special families of coefficients, explicit formulas for the density of zeros in a given set are available, see [5], [12], and the references therein. The following result states that the expected number of zeros in any compact set that does not meet $\mathbb{T}$ is of the order $O(\log n)$ as $n \to \infty$.

Proposition 3.4. *Let $E \subset \mathbb{C}$ be a compact set such that $E \cap \mathbb{T} = \emptyset$, and set $d := \operatorname{dist}(E, \mathbb{T})$. The expected number of zeros of P_n in E satisfies*

$$\mathbb{E}\left[n\tau_n(E)\right] \leq \frac{d+1}{d}\left(\frac{t+2}{t}\log(n+1) + \frac{2}{t}\log\mathbb{E}[|C_0|^t] - 2\mathbb{E}[\log|C_0|]\right). \tag{3.5}$$

If the set E does not have a "close" contact with the unit circle $\mathbb{T}$, then the expected number of zeros in E still remains of the order $o(n)$ as $n \to \infty$. In particular, we have the following

Proposition 3.5. *If E is a polygon inscribed in $\mathbb{T}$, then the expected number of zeros of P_n in E satisfies*

$$\mathbb{E}\left[n\tau_n(E)\right] = O(\sqrt{n\,\log n}) \quad \text{as } n \to \infty. \tag{3.6}$$

An estimate of this type for deterministic polynomials with restricted coefficients was proved in [6]. Another estimate for the zeros of deterministic polynomials in the disks $D_r(w) = \{z \in \mathbb{C} : |z-w| < r\},\ w \in \mathbb{T}$, is contained in [7]. We provide an analogue for the random polynomials below.

Proposition 3.6. *For $D_r(w)$ with $w \in \mathbb{T},\ r < 2$, the expected number of zeros of P_n in $D_r(w)$ satisfies*

$$\mathbb{E}\left[n\tau_n(D_r(w))\right] = \frac{2\arcsin(r/2)}{\pi}\, n + O\left(\sqrt{n\log n}\right) \quad \text{as } n \to \infty. \tag{3.7}$$

3.4. **Remarks on the non-iid case.** Under additional assumptions on ρ_C, we can say more. For instance, it is known (see [10, Chapter 4]) that if $\operatorname{Var}(|C_0|) = \sigma^2 < \infty$, then

$$\mathbb{E}[Y_n] \le \mu + \sigma \frac{n}{\sqrt{2n+1}},$$

and that equality can be achieved for each n with a particular choice of distribution. In this case, (3.3) takes on the asymptotic form $\mathbb{E}[\log Y_n^C] \le (1/2)\log n + O(1)$.

Moreover, the requirement that $C_0, C_1, \ldots$ be independent and identically distributed can be dropped if the second moments of their moduli are finite. Namely, if

$$\mathbb{E}[|C_0|] = \mathbb{E}[|C_1|] = \cdots = \mu \quad \text{and} \quad \operatorname{Var}(|C_0|) = \operatorname{Var}(|C_1|) = \cdots = \sigma^2, \tag{3.8}$$

then it follows from a result of Arnold and Groeneveld (see [10, Chapter 5]) that

$$\mathbb{E}[Y_n] \le \mu + \sigma\sqrt{n}. \tag{3.9}$$

Returning once more to (3.1), and applying Jensen's inequality, we deduce the following version of our result on expected discrepancies (which can then be applied to extend the other results in the previous subsection).

Theorem 3.7. *If the (not necessarily iid) coefficients of $P_n(z) = \sum_{k=0}^n C_k z^k$ have absolutely continuous distributions, and satisfy* (3.8)*, then*

$$\mathbb{E}\left[\left|\tau_n\left(A_r(\alpha,\beta)\right) - \frac{\beta-\alpha}{2\pi}\right|\right] \le \left(\sqrt{\frac{2\pi}{\mathbf{k}}} + \frac{2}{1-r}\right)\sqrt{\frac{\log(n+1) - \frac{1}{2}\mathbb{E}[\log|C_0|] - \frac{1}{2}\mathbb{E}[\log|C_n|] + O(1)}{n}},$$

as $n \to \infty$.

4. Examples

4.1. **Gaussian coefficients.** Let $C_0, C_1, \ldots$ be iid with density ρ^G given by

$$d\mu^G(z) = \rho^G(z)dA(z) = e^{-r^2}\frac{rdrd\theta}{\pi}; \tag{4.1}$$

that is, the coefficients are simply centered two-dimensional Gaussians. We readily compute that $\mathbb{E}[\log|C_0|] = -\gamma/2$.

We determine the asymptotics of $\mathbb{E}[\log(Y_n^G)]$ by elementary computations.

Lemma 4.1. *Let $n \geq 1$. Then*

$$\mathbb{E}[\log Y_n^G] = 2(n+1)\int_0^\infty x \log x\, e^{-x^2}(1-e^{-x^2})^n dx,$$

and, asymptotically,

$$\mathbb{E}[\log Y_n^G] = -\frac{\gamma}{2} + \frac{1}{2}\sum_{k=2}^{n+1}(-1)^k \binom{n+1}{k} \log k \leq \log\log n + O(1), \quad n \to \infty.$$

We obtain a slightly better expected discrepancy result (cf. [4, Sect. 3]).

Proposition 4.2. *For large enough n,*

$$\mathbb{E}\left[\left|\tau_n\left(A_r(\alpha,\beta)\right) - \frac{\beta-\alpha}{2\pi}\right|\right] \leq \left(\sqrt{\frac{2\pi}{\mathbf{k}}} + \frac{2}{1-r}\right)\sqrt{\frac{\log n + \log\log n + O(1)}{n}}.$$

4.2. **Heavy tails.** We now turn to the case of heavy-tailed coefficients.

Let $\alpha > 1$ and $C_0, C_1, \ldots$ be iid with a Pareto-type density

$$\rho^{P_\alpha}(z) = \begin{cases} \frac{\alpha-1}{2r^{\alpha+1}}, & |z| > 1, \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

Note that $\mathbb{E}[|C_0|] = (\alpha-1)/(\alpha-2)$ for $\alpha > 2$ and is infinite otherwise, and that $\mathbb{E}[\log|C_0|] = \mathbb{E}[\log^+|C_0|] < \infty$ for every $\alpha > 0$.

In view of (3.2), we obtain the following.

Lemma 4.3. *Let $n \geq 1$. Then*

$$\mathbb{E}[\log(Y_n^{P_\alpha})] = (\alpha-1)(n+1)\int_1^\infty \frac{\log x}{x^\alpha}(1-x^{-(\alpha-1)})^n dx = \frac{1}{\alpha-1}\mathbf{H}_{n+1},$$

where $\mathbf{H}_n = \sum_{k=1}^n 1/k$ is the nth harmonic number.

It is well-known that $\mathbf{H}_n = \log n + O(1)$, and so the estimate (3.3) cannot be improved in general.

5. Proofs

Let $D_r = \{z \in \mathbb{C} : |z| < r\},\ r > 0$. We need the following consequence of Jensen's formula.

Lemma 5.1. *If* $P_n(z) = \sum_{k=0}^n c_k z^k,\ c_k \in \mathbb{C}$ *and* $c_0 c_n \neq 0$, *then for any* $r \in (0,1)$ *we have*

$$\tau_n\left(\overline{D_r}\right) \le \frac{m(P_n) - \log|c_0|}{n(1-r)} \tag{5.1}$$

and

$$\tau_n\left(\mathbb{C} \setminus D_{1/r}\right) \le \frac{m(P_n) - \log|c_n|}{n(1-r)}. \tag{5.2}$$

Proof. Let $P_n(z) = c_n \prod_{k=1}^n (z - z_k)$. Using Jensen's formula, we obtain that

$$m(P_n) - \log|c_0| = \frac{1}{2\pi}\int_0^{2\pi} \log|P_n(e^{i\theta})|d\theta - \log|c_0| = \sum_{|z_k|<1} \log\frac{1}{|z_k|}$$
$$\ge \sum_{|z_k|\le r} \log\frac{1}{|z_k|} \ge n\tau_n\left(\overline{D_r}\right)\log\frac{1}{r} \ge n\tau_n\left(\overline{D_r}\right)(1-r).$$

Thus the first estimate follows, and we can apply it to the reciprocal polynomial $P_n^*(z) = z^n\overline{P_n(1/\bar{z})}$. Note that the zeros of P_n^* are $1/\bar{z}_k,\ k = 1, \ldots, n$, and its constant term is $\bar{c}_n$. Since $|P_n^*(z)| = |P_n(z)|$ for $|z| = 1$, we have that $m(P_n^*) = m(P_n)$. Hence (5.1) applied to P_n^* now gives that

$$\tau_n\left(\mathbb{C} \setminus D_{1/r}\right) \le \frac{m(P_n^*) - \log|\bar{c}_n|}{n(1-r)} = \frac{m(P_n) - \log|c_n|}{n(1-r)}.$$

□

Proof of Proposition 2.1. Under the assumptions of this proposition, the result of Mignotte [19] (see also [2]) gives

$$\left|\tau_n(S(\alpha,\beta)) - \frac{\beta-\alpha}{2\pi}\right| \le \sqrt{\frac{2\pi}{\mathbf{k}}}\sqrt{\frac{1}{n}\, m^+\left(\frac{P_n}{\sqrt{|c_0 c_n|}}\right)}.$$

On the other hand, applying Lemma 5.1, we obtain that

$$\tau_n(\mathbb{C} \setminus A_r(\alpha,\beta)) \le \frac{2}{n(1-r)}\, m\left(\frac{P_n}{\sqrt{|c_0 c_n|}}\right). \tag{5.3}$$

Since $\tau_n(A_r(\alpha,\beta)) = \tau_n(S(\alpha,\beta)) - \tau_n(\mathbb{C} \setminus A_r(\alpha,\beta))$, (2.1) follows as a combination of the above estimates. □

Proof of Proposition 2.2. Consider the set $E := \{\theta \in [0, 2\pi) : |P_n(e^{i\theta})| \ge 1\}$, and denote its length by $|E|$. Our assumption $\|P_n\|_p \ge 1$ implies that $|E| \neq 0$. We use concavity of log and Jensen's inequality in the following estimate

$$\begin{aligned} m^+(P_n) &= \frac{1}{2\pi}\int_0^{2\pi} \log^+ |P_n(e^{i\theta})|\, d\theta = \frac{1}{2\pi p}\int_E \log |P_n(e^{i\theta})|^p\, d\theta \\ &= \frac{|E|}{2\pi p}\int_E \log |P_n(e^{i\theta})|^p\, \frac{d\theta}{|E|} \le \frac{|E|}{2\pi p}\log\left(\int_E |P_n(e^{i\theta})|^p\, \frac{d\theta}{|E|}\right) \\ &= \frac{|E|}{2\pi p}\left(\log\frac{2\pi}{|E|} + \log\left(\frac{1}{2\pi}\int_E |P_n(e^{i\theta})|^p\, d\theta\right)\right) \\ &\le \frac{|E|}{2\pi p}\log\frac{2\pi}{|E|} + \frac{|E|}{2\pi p}\log\left(\frac{1}{2\pi}\int_0^{2\pi} |P_n(e^{i\theta})|^p\, d\theta\right) \\ &\le \frac{1}{p}\sup_{x\in(0,1]} x\log\frac{1}{x} + \log\|P_n\|_p = \frac{1}{ep} + \log\|P_n\|_p. \end{aligned}$$

□

Proof of Lemmas 3.1 and 3.2. In view of the fact that the C_k's are independent, we have

$$\begin{aligned} F_{Y_n}(r) &= \mathbb{P}(Y_n \le r) = \mathbb{P}(|C_0| \le r, |C_1| \le r, \ldots, |C_n| \le r) \\ &= \mathbb{P}(|C_0| \le r)\mathbb{P}(|C_1| \le r)\cdots\mathbb{P}(|C_n| \le r), \end{aligned}$$

and since they are identically distributed, $F_{Y_n}(r) = [\mathbb{P}(|C_0| \le r)]^{n+1}$. The statement of Lemma 3.1 now follows upon differentiation.

By Jensen's inequality,

$$\mathbb{E}[\log Y_n] \le \frac{1}{t}\log\mathbb{E}[Y_n^t],$$

and using (3.2), and $R_C(x) \le 1$, we obtain

$$\begin{aligned} \mathbb{E}[Y_n^t] = \int_0^\infty x^t (n+1)\rho_C(x)[R_C(x)]^n\, dx &\le (n+1)\int_0^\infty x^t\rho_C(x)dx \\ &= (n+1)\mathbb{E}[|C_0|^t], \end{aligned}$$

□

Proofs of Theorems 3.3 and 3.7. We first apply Proposition 2.1 and Jensen's inequality to obtain

$$\begin{aligned} \mathbb{E}\left[\left|\tau_n\left(A_r(\alpha,\beta)\right) - \frac{\beta-\alpha}{2\pi}\right|\right] \le \sqrt{\frac{2\pi}{\mathbf{k}}}\sqrt{\frac{1}{n}\,\mathbb{E}\left[m^+\left(\frac{P_n}{\sqrt{|C_0C_n|}}\right)\right]} \\ + \frac{2}{n(1-r)}\,\mathbb{E}\left[m\left(\frac{P_n}{\sqrt{|C_0C_n|}}\right)\right]. \end{aligned}$$

A combination of Proposition 2.2 with (3.1) and Lemma 3.2 gives that

$$\begin{aligned}\mathbb{E}\left[m^+\left(\frac{P_n}{\sqrt{|C_0C_n|}}\right)\right] &\le \mathbb{E}\left[\log\left\|\frac{P_n}{\sqrt{|C_0C_n|}}\right\|_2\right]+\frac{1}{2e}\\ &\le \frac{\log(n+1)}{2}+\mathbb{E}\left[\log\max_{k=0,\dots,n}|C_k|\right]-\mathbb{E}[\log|C_0|]+\frac{1}{2e}\\ &\le \frac{t+2}{2t}\log(n+1)+\frac{1}{t}\log\mathbb{E}[|C_0|^t]-\mathbb{E}[\log|C_0|]+\frac{1}{2e}. \qquad (5.4)\end{aligned}$$

To justify the use of Proposition 2.2, we note that $\left\|P_n/\sqrt{|C_0C_n|}\right\|_2 \ge 1$, which is a consequence of the fact that $\|P_n\|_2 \ge |C_k|,\ k=0,\dots,n$.

Using (1.4), (3.1) and Lemma 3.2, we obtain that

$$\begin{aligned}\mathbb{E}\left[m\left(\frac{P_n}{\sqrt{|C_0C_n|}}\right)\right] &\le \mathbb{E}\left[\log\left\|\frac{P_n}{\sqrt{|C_0C_n|}}\right\|_2\right]\\ &\le \frac{\log(n+1)}{2}+\mathbb{E}\left[\log\max_{k=0,\dots,n}|C_k|\right]-\mathbb{E}[\log|C_0|]\\ &\le \frac{t+2}{2t}\log(n+1)+\frac{1}{t}\log\mathbb{E}[|C_0|^t]-\mathbb{E}[\log|C_0|].\end{aligned}$$

Thus (3.4) follows from the above estimates.

The proof of Theorem 3.7 is similar. We first argue as above, bounding the discrepancy in terms of $\mathbb{E}\left[\log\left\|P_n/\sqrt{|C_0C_n|}\right\|_2\right]$. We then have

$$\begin{aligned}\mathbb{E}\left[\log\left\|P_n/\sqrt{|C_0C_n|}\right\|_2\right] \le \frac{\log(n+1)}{2}&+\mathbb{E}[\log(\max_k|C_k|)]\\ &-\frac{\mathbb{E}[\log|C_0|]+\mathbb{E}[\log|C_n|]}{2}.\end{aligned}$$

We now appeal to (3.9) and Jensen's inequality instead of Lemma 3.2 to obtain

$$\mathbb{E}[\log(\max_k|C_k|)] \le \log[\mu+\sigma\sqrt{n}] \le \frac{1}{2}\log(n+1)+O(1),$$

and the theorem follows. □

Proof of Proposition 3.4. Using Lemma 5.1, we obtain as in (5.3) that

$$\tau_n\left(\mathbb{C}\setminus A_r(0,2\pi)\right) \le \frac{2}{n(1-r)}\,m\left(\frac{P_n}{\sqrt{|C_0C_n|}}\right).$$

If r is selected so that $E\subset\mathbb{C}\setminus A_r(0,2\pi)$, then

$$\begin{aligned}\mathbb{E}\left[n\tau_n(E)\right] &\le \frac{2}{1-r}\,\mathbb{E}\left[m\left(\frac{P_n}{\sqrt{|C_0C_n|}}\right)\right] \le \frac{2}{1-r}\,\mathbb{E}\left[\log\left\|\frac{P_n}{\sqrt{|C_0C_n|}}\right\|_2\right]\\ &\le \frac{2}{1-r}\left(\frac{t+2}{2t}\log(n+1)+\frac{1}{t}\log\mathbb{E}[|C_0|^t]-\mathbb{E}[\log|C_0|]\right),\end{aligned}$$

where we used (1.4), (3.1) and Lemma 3.2. An elementary argument shows that $r = 1/(\operatorname{dist}(E, \mathbb{T}) + 1)$ implies $E \subset \mathbb{C} \setminus A_r(0, 2\pi)$. □

Proof of Proposition 3.5. Suppose that the vertices of our polygon are located at the points $e^{i\theta_j}$, $j = 1, \ldots, k$. It follows from a simple Euclidean geometry consideration that the polygon is contained in the union of the closed disk $U = \{z \in \mathbb{C} : |z| \leq 1 - \sqrt{\log n/n}\}$ and sectors $S_j = \{z \in \mathbb{C} : |\arg z - \theta_j| < C\sqrt{\log n/n}\}$, for each $n \in \mathbb{N}$, where $C > 0$ is a constant that depends only on the polygon. Note that Proposition 3.4 applies to U, so that (3.5) gives

$$\mathbb{E}\left[n\tau_n(U)\right] = O\left(\sqrt{n \log n}\right) \quad \text{as } n \to \infty. \tag{5.5}$$

We recall the result of Mignotte [19] (see also [2])

$$\left|\tau_n\left(S(\alpha, \beta)\right) - \frac{\beta - \alpha}{2\pi}\right| \leq \sqrt{\frac{2\pi}{\mathbf{k}}} \sqrt{\frac{1}{n} m^+\left(\frac{P_n}{\sqrt{|C_0 C_n|}}\right)},$$

and apply it to each sector S_j. It follows from the above estimate and Jensen's inequality that

$$\mathbb{E}\left[n\tau_n(S_j)\right] = O\left(\sqrt{n \log n}\right) + O\left(\sqrt{n\mathbb{E}\left[m^+\left(\frac{P_n}{\sqrt{|C_0 C_n|}}\right)\right]}\right) \quad \text{as } n \to \infty,$$

for each $j = 1, \ldots, k$. Using the already available inequality (5.4), we conclude that the second term is of the same order as the first, so that

$$\mathbb{E}\left[n\tau_n(S_j)\right] = O\left(\sqrt{n \log n}\right) \quad \text{as } n \to \infty,$$

for each $j = 1, \ldots, k$. Thus (3.6) is a consequence of (5.5) and the above equation. □

Proof of Proposition 3.6. We follow ideas similar to those used in the proof of Proposition 3.5. Note that the intersection of the disk $D_r(w)$ and $\mathbb{T}$ is an arc with endpoints $e^{i\alpha}$ and $e^{i\beta}$, where $\alpha < \beta < \alpha + 2\pi$ and $\beta - \alpha = 4\arcsin(r/2)$. Furthermore, $D_r(w)$ is contained in the union of the circular sector $S = \{z \in \mathbb{C} : \alpha - C\sqrt{\log n/n} < \arg z < \beta + C\sqrt{\log n/n}\}$, where $C > 0$ depends only on r, and the set $F = \{z \in \mathbb{C} : |z| \leq 1 - \sqrt{\log n/n} \text{ or } |z| \geq 1 + \sqrt{\log n/n}\}$. Proposition 3.4 implies that

$$\mathbb{E}\left[n\tau_n(F)\right] = O\left(\sqrt{n \log n}\right) \quad \text{as } n \to \infty.$$

On the other hand, Mignotte's estimate gives that

$$\mathbb{E}\left[n\tau_n\left(S\right) - \frac{\beta - \alpha}{2\pi} n\right] = O\left(\sqrt{n \log n}\right) \quad \text{as } n \to \infty,$$

arguing as in the proof of Proposition 3.5. Combining the last two equations, we arrive at (3.7). □

Proofs of Lemmas 4.1 and 4.3. The first statement follows from Lemma 3.1 and the fact that $R^G(x) = \int_0^x r\, e^{-r^2} dr = (1 - e^{-x^2})/2$; the series expression is readily obtained by using the binomial theorem.

To analyze the asymptotics of $\mathbb{E}[\log Y_n^G]$, we first perform the obvious change of variables $u = \exp(-x^2)$ to obtain

$$\begin{aligned}\mathbb{E}[\log Y_n^G] &= \frac{n+1}{2} \int_0^1 \log\log\left(\frac{1}{u}\right)(1-u)^n\, du \\ &= I_1(n) + I_2(n) = \int_0^{\frac{1}{(n+1)^2}} + \int_{\frac{1}{(n+1)^2}}^1 .\end{aligned}$$

Since $\log\log(1/u)$ is decreasing, for all sufficiently large n, the latter integral admits the estimate

$$\begin{aligned}I_2(n) &\le \frac{(n+1)\log\log(n+1)^2}{2} \int_{\frac{1}{(n+1)^2}}^1 (1-u)^n du \\ &= \frac{\log\log(n+1)^2}{2}\left(1 - \frac{1}{(n+1)^2}\right)^{n+1} \le \frac{\log\log(n+1)^2}{2}.\end{aligned}$$

We next show that the contribution arising from the first integral is negligible. An integration by parts shows that

$$\int_0^x \log\log\left(\frac{1}{u}\right) du = x\log\log\left(\frac{1}{x}\right) - \mathrm{li}(x), \quad x > 0,$$

where $\mathrm{li}(x) = \int_0^x 1/\log(u)du$ is the logarithmic integral. We use this identity together with the crude estimate $(1-u)^n \le 1$ to obtain

$$\begin{aligned}I_1(n) &\le \frac{n+1}{2} \int_0^{\frac{1}{(n+1)^2}} \log\log\left(\frac{1}{u}\right) du \\ &= \frac{1}{2(n+1)} \log\log(n+1)^2 - \frac{n+1}{2}\mathrm{li}\left(\frac{1}{(n+1)^2}\right).\end{aligned} \tag{5.6}$$

The first term on the right-hand side in (5.6) clearly tends to zero as $n \to \infty$. We then note that the function $x \mapsto \mathrm{li}(x^2)/x$ is continuous on $(0,1)$, and by L'Hôpital's rule, $\lim_{x\to 0} \mathrm{li}(x^2)/x = \lim_{x \to 0} x/\log(x) = 0$. Hence $I_1(n) = o(1)$ as required.

For ρ^{P_α}, we have

$$\begin{aligned}\mathbb{E}[\log(Y_n^{P_\alpha})] &= (n+1)\int_0^1 \log\left[\frac{1}{(1-u)^{\frac{1}{\alpha-1}}}\right] u^n\, du \\ &= \frac{n+1}{\alpha-1}\int_0^1 \log\frac{1}{1-u}\, u^n\, du,\end{aligned}$$

and the result follows from Euler's formula $\mathbf{H}_n = \int_0^1 \frac{1-u^n}{1-u} du$ and an integration by parts. □

Acknowledgments

The second author thanks Martin Bender for interesting discussions.

References

[1] R. J. Adler, The geometry of random fields, John Wiley & Sons, Chichester, 1981.

[2] F. Amoroso and M. Mignotte, On the distribution of the roots of polynomials, Annales Inst. Fourier (Grenoble) **45** (1996), 1275–1291.

[3] V. V. Andrievskii and H.-P. Blatt, Discrepancy of signed measures and polynomial approximation, Springer Monogr. Math., Springer-Verlag, New York, 2002.

[4] L. Arnold, Über die Nullstellenverteilung zufälliger Polynome, Math. Z. **92** (1966), 12-18.

[5] A. T. Bharucha-Reid and M. Sambandham, Random polynomials, Academic Press, Orlando, 1986.

[6] P. Borwein and T. Erdélyi, On the zeros of polynomials with restricted coefficients, Illinois J. Math. **41** (1997), 667–675.

[7] P. Borwein, T. Erdélyi and F. Littmann, Polynomials with coefficients from a finite set, Trans. Amer. Math. Soc. **360** (2008), 5145–5154.

[8] P. Borwein and R. Lockhart, The expected L_p norm of random polynomials, Proc. Amer. Math. Soc. **129** (2000), 1463–1472.

[9] K.-K. S. Choi and M. J. Mossinghoff, Average Mahler's measure and L_p norms of unimodular polynomials, Pacific J. Math. **252** (2011), 31-50.

[10] H. A. David and H. N. Nagaraja, Order statistics, Third ed., Wiley Series in Probability and Statistics, John Wiley & Sons, Hoboken, NJ (2003).

[11] P. Erdős and P. Turán, On the distribution of roots of polynomials, Annals of Math. **51** (1950), 105–119.

[12] K. Farahmand, Topics in random polynomials, Pitman Res. Notes Math. **393** (1998).

[13] G. T. Fielding, The expected value of the integral around the unit circle of a certain class of polynomials, Bull. London Math. Soc. **2** (1970), 301–306.

[14] T. Ganelius, Sequences of analytic functions and their zeros, Ark.Mat.**3** (1958), 1–50.

[15] C. P. Hughes and A. Nikeghbali, The zeros of random polynomials cluster uniformly near the the unit circle, Compositio Math. **144** (2008), 734–746.

[16] I. Ibragimov and O. Zeitouni, On roots of random polynomials, Trans. Amer. Math. Soc. **349** (1997), 2427–2441.

[17] I. Ibragimov and D. Zaporozhets, On distribution of zeros of random polynomials in complex plane, available at http://arxiv.org/abs/1102.3517.

[18] Z. Kabluchko and D. Zaporozhets, Roots of random polynomials whose coefficients have logarithmic tails, available at http://arxiv.org/abs/1110.2585 (to appear in Annals of Probab.).

[19] M. Mignotte, Remarque sur une question relative à des fonctions conjuguées, C. R. Acad. Sci. Paris Ser. I **315** (1992), 907–911.

[20] I. E. Pritsker, Means of algebraic numbers in the unit disk, C. R. Acad. Sci. Paris Ser. I **347** (2009), 119–122.

[21] I. E. Pritsker, Equidistribution of points via energy, Ark. Mat. **49** (2011), 149–173.

[22] E. Shmerling and K.J. Hochberg, Asymptotic behavior of roots of random polynomial equations, Proc. Amer. Math. Soc. **130** (2002), 2761-2770.

Department of Mathematics, Oklahoma State University, Stillwater, OK 74078, USA.

E-mail address: igor@math.okstate.edu

Statistical Laboratory, University of Cambridge, Cambridge CB3 0WB,UK.

E-mail address: a.sola@statslab.cam.ac.uk