REAL ROOTS OF RANDOM POLYNOMIALS: EXPECTATION AND REPULSION

YEN DO, HOI NGUYEN, AND VAN VU

ABSTRACT. Let $P_n(x) = \sum_{i=0}^n \xi_i x^i$ be a Kac random polynomial where the coefficients ξ_i are iid copies of a given random variable ξ. Our main result is an optimal quantitative bound concerning real roots repulsion. This leads to an optimal bound on the probability that there is a real double root.

As an application, we consider the problem of estimating the number of real roots of P_n, which has a long history and in particular was the main subject of a celebrated series of papers by Littlewood and Offord from the 1940s. We show, for a large and natural family of atom variables ξ, that the expected number of real roots of $P_n(x)$ is exactly $\frac{2}{\pi}\log n + C + o(1)$, where C is an absolute constant depending on the atom variable ξ. Prior to this paper, such a result was known only for the case when ξ is Gaussian.

1. INTRODUCTION

Let ξ be a real random variable having no atom at 0, zero mean and unit variance. Our object of study is the random polynomial

$$P_n(x) := \sum_{i=0}^n \xi_i x^i \tag{1}$$

where ξ_i are iid copies of ξ. This polynomial is often referred to as Kac's polynomial, and has been extensively investigated in the literature.

1.1. Real roots of random polynomials.

The study of real roots of random polynomials has a long history. Let N_n be the number of real roots of $P_n(x)$, sometimes we use the notation $N_{n,\xi}$ to emphasize the dependence of N_n on the distribution of ξ. This is a random variable taking values in $\{0, \dots, n\}$. The issue of estimating N_n was already raised by Waring as far back as 1782 ([32, page 618], [16]), and has generated a large amount of literature, of which we give an (incomplete and brief) survey.

One of the first estimates for N_n was obtained by Bloch and Pólya [1], who studied the case when ξ is uniformly distributed in $\{-1, 0, 1\}$, and established the upper bound

$$\mathbf{E}N_n = O(n^{1/2}).$$

Here and later we'll use the usual asymptotic notation $X = O(Y)$ or $X \ll Y$ to denote the bound $|X| \leq CY$ where C is independent of Y. The above bound of Bloch and Pólya was not sharp, as it

2010 *Mathematics Subject Classification.* 15A52, 60C05, 60G50.

Y. Do is supported in part by research grant DMS-1201456.

H. Nguyen is supported by research grant DMS-1358648.

V. Vu is supported by NSF Grant DMS-1307797 and AFORS Grant FA9550-12-1-0083. Part of this work was done at VIASM (Hanoi), and the author would like to thank their support and hospitality.

turned out later that P_n has a remarkably small number of real roots. In a series of breakthrough papers [17, 18, 19, 20] in the early 1940s, Littlewood and Offord proved (for many atom variables ξ such as Gaussian, Bernoulli or uniform on $[-1,1]$) that

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

with probability $1-o(1)$, where we use $o(1)$ to denote a quantity that goes to 0 as $n \to \infty$.

Around the same time, Kac [13] developed a general formula for the expectation of number of real roots

$$\mathbf{E}N_n = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} |y| p(t,0,y) dy, \tag{2}$$

where $p(t,x,y)$ is the probability density for $P_n(t) = x$ ad $P_n'(t) = y$.

In the Gaussian case, one can easily evaluate the RHS and get

$$\mathbf{E}N_n = \frac{1}{\pi}\int_{-\infty}^{\infty} \sqrt{\frac{1}{(t^2-1)^2} + \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}} dt = (\frac{2}{\pi} + o(1)) \log n. \tag{3}$$

For non-Gaussian distributions, however, Kac's formula is often very hard to evaluate and it took a considerable amount of work to extend (3) to other distributions. In a subsequent paper [14], Kac himself handled the case when ξ is uniformly distributed on the interval $[-1,1]$ and Stevens [28] extended it further to cover a large class of ξ having continuous and smooth distributions with certain regularity properties (see [28, page 457] for details). These papers relies on (2) and the analytic properties of the distribution of ξ.

For discrete distributions, (2) does not appear useful and it took more than 10 years since Kac's paper until Erdős and Offord in 1956 [7] found a completely new approach to handle the Bernoulli case. For this case, they proved that with probability $1 - o(\frac{1}{\sqrt{\log\log n}})$

$$N_{n,\xi} = \frac{2}{\pi}\log n + o(\log^{2/3} n \log\log n). \tag{4}$$

In the late 1960s and early 1970s, Ibragimov and Maslova [9, 10] successfully refined Erdős-Offord's method to handle any variable ξ with mean 0. They proved that for any ξ with mean zero which belong to the domain of attraction of the normal law,

$$\mathbf{E}N_{n,\xi} = \frac{2}{\pi}\log n + o(\log n). \tag{5}$$

For related results, see also [11, 12]. Few years later, Maslova [23, 24] showed that if ξ has mean zero and variance one and $\mathbf{P}(\xi = 0) = 0$, then the variance of $N_{n,\xi}$ is $(\frac{4}{\pi}(1-\frac{2}{\pi}) + o(1))\log n$.

Other developments were made in the late 1980s by Wilkins [34] and in the early 1990s by Edelman and Kostlan [5], who evaluated the explicit integral in (3) very carefully and provided a precise estimate for $\mathbf{E}N_{n,N(0,1)}$

$$\mathbf{E}N_{n,N(0,1)} = \frac{2}{\pi}\log n + C_{Gau} + o(1). \tag{6}$$

where $C_{Gau} \approx .625738072..$ is an explicit constant (the value of an explicit, but complicated integral). As a matter of fact, one can even write $o(1)$ as sum of explicit functions of n, which gives a complete Taylor expansion.

The truly remarkable fact about (6) is that the error term $\mathbf{E}N_{n,N(0,1)} - \frac{2}{\pi}\log n$ tends to a limit as n tends to infinity. The question here is: Is this a *universal phenomenon*, which holds for general random polynomials, or a special property of the Gaussian one ?

It is clear that the computation leading to (6) is not applicable for general random polynomials, as the explicit formula in (3) is available only in the Gaussian case, thanks to the unitary invariance property of this particular distribution. For many natural variables, such as Bernoulli, there is little hope that such an explicit formula actually exists. As a matter of fact, Ibragimov-Maslova's proof of their asymptotic result for general non-Gaussian polynomials (based on the earlier work of Erdős-Offord) is a tour-de-force. Among others, they followed the Erdős-Offord's idea of using the number of sign changes on a fixed sequence of points to approximate the number of roots. The error term in this approximation, by nature, has to be large (at least a positive power of $\log n$).

On the other hand, numerical evidence tends to support the conjecture that $\mathbf{E}N_n - \frac{2}{\pi}\log n$ do go to a limit, as n to tends to infinity. However, the situation is delicate as this limit seems to depend on the distribution of the atom variable ξ and *is not* universal; see the numerical illustration below.

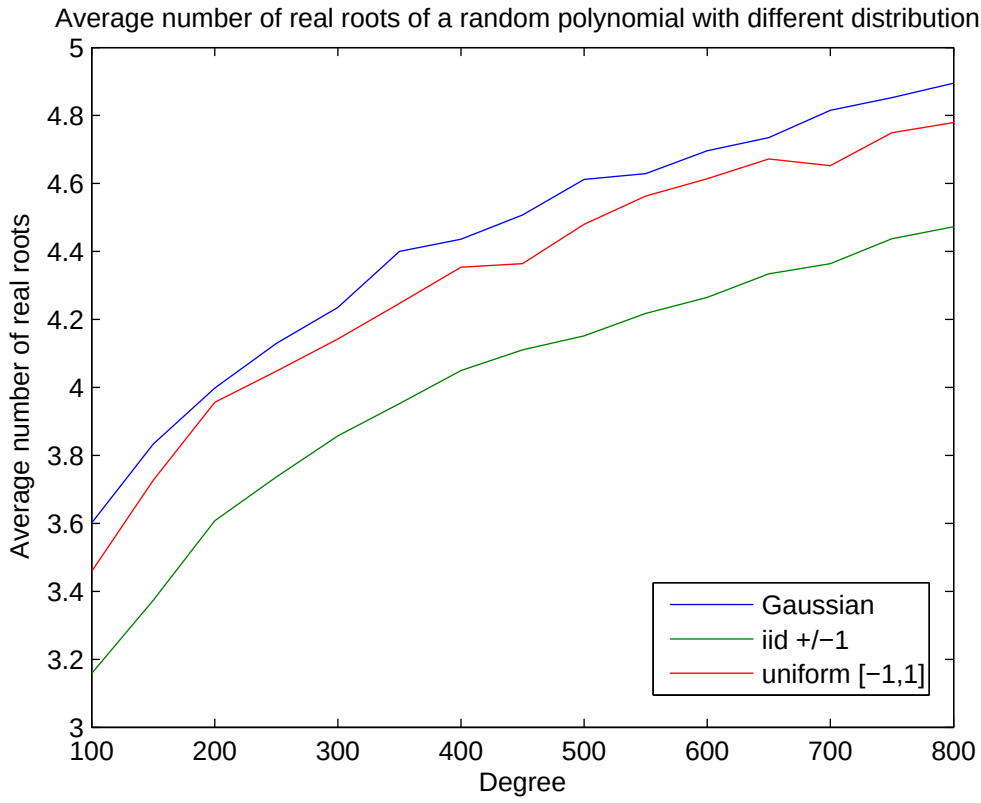

FIGURE 1. The figure shows average numbers of real roots of random polynomials with different distributions. One can see the shape of the curve $\frac{2}{\pi}\log n$ in all three cases.

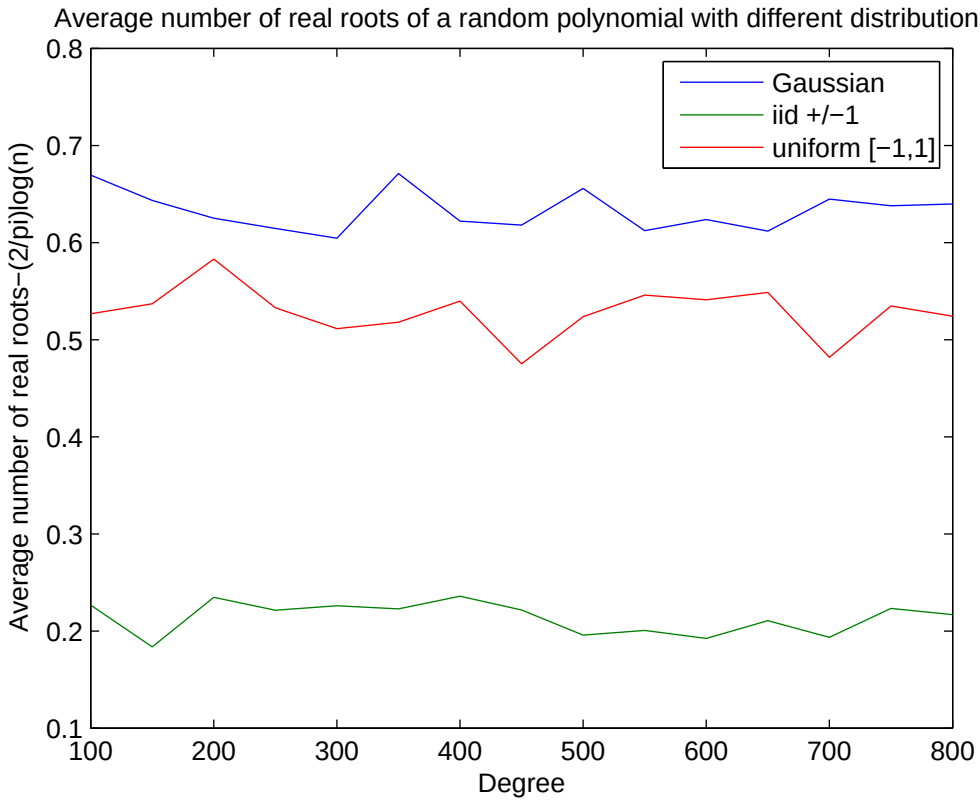

FIGURE 2. In this second figure, we subtract $\frac{2}{\pi}\log n$ from the averages, and the curves seem to converge to different values.

In a recent work [26], the last two authors and Oanh Nguyen made a first step by showing that the error term in question is bounded.

Theorem 1.2. *Let ξ be a random variable with mean 0 and variance 1 and bounded $(2+\epsilon)$-moment. Then*

$$|\mathbf{E}N_{n,\xi} - \frac{2}{\pi}\log n| = O_{\epsilon,\xi}(1). \tag{7}$$

The approach in [26], however, reveals little about the behavior of the bounded quantity $O_{\epsilon,\xi}(1)$.

In this paper, we settle this problem in the affirmative for a large and natural class of distributions, as an application of a general theorem concerning the repulsion between real roots of Kac's polynomials (see the next section).

Definition 1.3 (Type I, discrete distributions)**.** For any positive integer N, we say that ξ has uniform distribution with parameter N (or *type I*) if $\mathbf{P}(\xi = i) = 1/(2N)$ independently, $i \in \{\pm 1, \pm 2, \dots, \pm N\}$. In particularly, Bernoulli random variable has uniform distribution with parameter 1.

Definition 1.4 (Type II, continuous distributions)**.** Let $\varepsilon_0 > 0$ and $p > 1$. We say that a random variable ξ of mean zero has *type II* distribution with parameter (p, ε_0) if its has a p-integrable density function and its $(2+\varepsilon_0)$-moment is bounded.

Theorem 1.5 (Main application: Expectation of real roots)**.** *Let ξ be a random variable with either type I or type II with fixed parameters. Then*

$$\mathbf{E}N_{n,\xi} = \frac{2}{\pi}\log n + C + o(1),$$

where C is an absolute constant depending on ξ.

We would like to mention that due to the abstract nature of our arguments, the exact value of C is still unknown, and its determination remains a tantalizing question. In our favorite toy case when ξ is Bernoulli, computer simulation suggests that C is around .22, but it looks already difficult to prove that $|C| \leq 10$ (say).

Now we are going to present the main technical contribution of our paper, which, together with some other tools, will yield Theorem 1.5 as an application. The object of study here is the existence of double roots, or more generally, the repulsion between real roots of Kac polynomials.

1.6. **Roots repulsion of random polynomials.** Multiple roots is a basic object in theoretical analysis. They also play an important role in practical computational problems. For example, it is a fundamental fact in numerical analysis that the running time of Newton's method (for finding real roots of a function) increases exponentially with the presence of a multiple or near multiple root (i.e., a place x where both $|P(x)|$ and $|P'(x)|$ are close to zero), see for instance [2, Chapter 8].

Intuitively, one expects that a random polynomial does not have multiple or near multiple roots, with high probability. Motivated by problems in complexity theory, researchers have confirmed this intuition for the case when ξ is Gaussian, in a more general setting (see [2, Chapter 13] and [4]). Unfortunately, the methods in these works rely on the invariance property of the gaussian distribution and do not extend to other natural distributions.

We are going to introduce a new approach that enables us to fully understand the double root phenomenon, and consequently derive an optimal bound on the probability that a random polynomial has double or near double real roots.

For the sake of presentation, let us first consider the toy case when ξ is Bernoulli. One expects that the probability of having double roots tends to 0 with n; the question is how fast ? To give a lower bound, let us consider $p_{\pm 1}$, the probability that the polynomial has a double root at either 1 or -1 (we stress that $p_{\pm 1}$ denotes one number). Trivially, the double root probability is at least $p_{\pm 1}$. A short consideration shows (see Appendix A)

Fact 1.7. *Assume that ξ_j's have the Bernoulli distribution. Then $p_{\pm 1} = \Theta(n^{-2})$ if $4|(n+1)$ and 0 otherwise.*

We are now ready to state our first theorem, which asserts that in the $4|(n+1)$ case, the double root probability is dominated by $p_{\pm 1}$, while in the other cases, this probability is very small.

Theorem 1.8 (Double roots of Bernoully polynomials)**.** *Let $P_n(x)$ be a Bernoulli polynomial. Then*

$$\mathbf{P}(P_n \text{ has real double roots}) = p_{\pm 1} + n^{-\omega(1)}.$$

Theorem 1.8 is a special case of our main result below, which deals with near double roots of general random polynomials.

Theorem 1.9 (Main result: Roots repulsion)**.** *Assume that $\xi_0, \dots, \xi_n$ are independent and all of them are either Type I with the same fixed parameter N, or Type II with uniform implicit constants (which are fixed). Then for any constant $C > 0$ there exists a constant $B > 0$ such that for all sufficiently large n*

$$\mathbf{P}\Big(\exists x \in \mathbb{R} : P_n(x) = 0 \wedge |P_n'(x)| \le n^{-B}\Big) \le p_{\pm 1} + n^{-C}$$

where $p_{\pm 1}$ denotes the probability that the polynomial has a double root at either 1 or -1.

It is clear that in the type II setting we have $p_{\pm 1} = 0$. Also, it is not hard to show that $p_{\pm 1} = O(1/n^2)$ for the type I setting (see Claim A.5.)

Our proof will provide more detailed information about the location of double and near double roots. For instance, in the discrete case, we will show that with overwhelming probability, any double (or near double) root, has to be very close to either 1 or -1. We present the precise statements in Section 2 and with proofs in Sections 3-5. As an application, we verify Theorem 1.5 in Section 6. A few technical statements will be proved in the Appendices in order to maintain the flow of the presentation.

Remark 1.10. *Parallel to this paper, Peled et. al.* [27] *proved for random polynomials with the atom variable ξ having support on $\{0, \pm 1\}$ that the probability of having double roots (both real and complex) is dominated by the probability of having double roots at $0, \pm 1$. The method used in* [27] *is specialized for the support consisting of $\{0, \pm 1\}$ and is totally different from the method we used in this paper.*

2. More precise statements

The harder case in our study is the discrete one (ξ is of Type I). In what follows we first discuss our approach for this case.

2.1. **ξ of type I.** Note that the real roots of $P_n(x)$ have absolute value bounded from above by $N + 1$ and below by $1/(N + 1)$. It follows that we only need to consider $1/(N + 1) < |x| < N + 1$. Since $P_n(-x)$ and $x^n P_n(1/x)$ have the same distribution as $P_n(x)$, it suffices to consider $1/(N + 1) < x \le 1$.

Let $0 < \varepsilon < 2$ be a constant and set

$$I_0 := (1/(N + 1), 1 - n^{-2+\varepsilon}], \text{ and } I_1 := (1 - n^{-2+\varepsilon}, 1]. \tag{8}$$

Theorem 1.9 is a consequence of the following two results, where ξ is assumed to have type I.

Theorem 2.2 (No double root in the bulk)**.** *Let $C > 0$ be any constant. Then there exists $B > 0$ depending on C and ε such that the following holds*

$$\mathbf{P}\Big(\exists x \in I_0 : |P_n(x)| \le n^{-B}, |P_n'(x)| \le n^{-B}\Big) = O(n^{-C}).$$

Theorem 2.3 (No double root at the edge)**.** *For sufficiently small ε, there exists $B > 0$ such that*

$$\mathbf{P}\Big(\exists x \in I_1 : |P_n(x)| \le n^{-B}, |P_n'(x)| \le n^{-B}\Big) = \mathbf{P}\Big(P_n(1) = P_n'(1) = 0\Big).$$

In fact, it follows from the proof in Section 5 that one can take $\varepsilon \le 1/8$ and $B = 16$.

To prove Theorem 1.5, we will need the following stronger version of Theorem 2.2.

Theorem 2.4. *Assume that ξ has uniform discrete distribution with parameter N, where N is fixed. Let $C > 0$ be any constant. Then there exists $B > 0$ depending on C, N and ε such that the following holds with probability at least $1 - O(n^{-C})$.*

(i) (Near double roots) There does not exist $x \in I_0$ such that
$$|P_n(x)| \le n^{-B} \wedge |P_n'(x)| \le n^{-B}.$$
(ii) (Repulsion) There do not exist $x, x' \in I_0$ with $|x - x'| \le n^{-B}$ such that
$$P_n(x) = P_n(x') = 0.$$
(iii) (Delocalization) For any given $a \in I_0$, there is no x with $|P_n(x)| \le n^{-B}$ and $|x - a| \le n^{-B}$.

We remark that Theorem 2.4 might continue to hold for an interval larger than I_0, but we do not try to pursue this matter here. We next turn to the continuous case.

2.5. **ξ of type II.** For any interval $I \subset \mathbb{R}$, using Hölder's inequality we have
$$P(\xi \in I) = O(|I|^{1-1/p}) \quad .$$

Since $p > 1$, it follows that $P(|\xi| < n^{-C}) = O(n^{-C(p-1)/p})$. Additionally, as ξ has bounded $(2+\varepsilon_0)$-moment, we have $\mathbf{P}(|\xi| > n^C) = \mathbf{P}(|\xi|^{2+\varepsilon_0} > n^{(2+\varepsilon_0)C}) = O(n^{-(2+\varepsilon_0)C})$. Therefore with a loss of at most $O(n^{-C})(C > 1)$ in probability one can assume that

$$n^{-C_1} \le |\xi_i| \le n^{C_1}, \forall 1 \le i \le n \quad , \tag{9}$$

where C_1 is a finite constant depending on p and C.

Conditioning on this, it can be shown easily that if $|x| \le \frac{1}{4} n^{-2C_1}$ then $|P_n(x)| \ge n^{-C_1}/2$ and if $|x| \ge 4n^{2C_1}$ then $|P_n(x)| \ge n^{-C_1} x^n/2 \gg 1$. It follows that Theorem 1.9 follows from the following analogue of Theorem 2.4. We remark that in this theorem, we allow a more general setting where the coefficients ξ_i are not necessarily iid, which is convenient in the proof.

Theorem 2.6. *Assume that $\xi_0, \dots, \xi_n$ have type II distributions with uniform implicit constants. Consider $P_n(x) = \xi_n x^n + \dots + \xi_0$. Let $C > 1$ be any constant. Then there exists $B > 0$ depending on C, ε_0 and p such that (i),(ii),(iii) of Theorem 2.4 hold with probability at least $1 - O(n^{-C})$ with I_0 replaced by $I_0' := [-4n^{2C_1}, -\frac{1}{4} n^{-2C_1}] \cup [\frac{1}{4} n^{-2C_1}, 4n^{2C_1}]$. In other words,*

(i) there does not exist $x \in I_0'$ such that
$$|P_n(x)| \le n^{-B} \wedge |P_n'(x)| \le n^{-B};$$
(ii) there do not exist $x, x' \in I_0'$ with $|x - x'| \le n^{-B}$ such that
$$P_n(x) = P_n(x') = 0;$$
(iii) for any given $a \in I_0'$, there is no x with $|P_n(x)| \le n^{-B}$ and $|x - a| \le n^{-B}$.

In the next section, we discuss the strategy to prove Theorems 2.4 and 2.6.

3. The general strategy and the proof of Theorem 2.4

In this section, we first explain our strategy to prove (the harder) Theorem 2.4 and then deduce Theorem 2.6 from this approach. The rest of the proof of Theorem 2.4 follows in the next section.

Our general strategy is to reduce the event of having double (or near double) roots to the event that a certain random variable takes value in a small interval. The key step is to bound the probability of the latter, and here our main tool will be a recently developed machinery, the so-called Inverse Littlewood-Offord theory (see [25] for an introduction).

Divide the interval I_0 into subintervals of length δ each (except for possibly the right-most interval which has the same right endpoint as I_0), where δ to be chosen sufficiently small (polynomially in n), and B is then chosen large enough so that $\delta^2 \gg n^{-B}$.

Near double roots. Assume that there exists a subinterval I and an element $x \in I$ such that $|P_n(x)| \leq n^{-B}$ and $|P_n'(x)| \leq n^{-B}$, then for x_I, the center of I, we have

$$|P_n(x_I)| \leq \delta |P_n'(y)| + n^{-B}$$

for some $y \in I$.

In the following, the implicit constants in $O(.)$ may depend on N unless otherwise specified.

On the other hand, as $|P_n'(y)| \leq \delta |P_n''(z)| + n^{-B}$ for some $z \in I$. From here, by the trivial upper bound $O(n^3)$ for the second derivative, we have

$$|P_n(x_I)| = O(\delta^2 n^3 + n^{-B}). \tag{10}$$

Repulsion. Assume that $P_n(x) = P_n(x') = 0$ for some $x, x' \in I_0$ with $|x - x'| \leq n^{-B}$. Then there is a point y between x and x' such that $P_n'(y) = 0$. Thus, for any z with$|z - y| \leq 2\delta$,

$$|P'(z)| \leq 2\delta n^3.$$

There is a point x_I of some subinterval I such that $|x_I - x| \leq \delta$. For this x_I, $|P_n(x_I)| = |x_I - x||P_n'(z)|$ for some z between x and x_I. Because x has distance at most $n^{-B} \ll \delta$ from x', x also has distance at most δ from y, and so z has distance at most 2δ from y. It follows that

$$|P_n(x_I)| \leq 2\delta^2 n^3. \tag{11}$$

Remark 3.1. *One can also show that the* **repulsion** *property is a direct consequence of the* **near double roots** *property by choosing B slightly larger if needed. Indeed, suppose that $P_n(x) = P_n(x') = 0$ for some $x, x' \in I_0$ with $|x - x'| \leq n^{-B}$, then consider the y obtained as above. Thus $P'(y) = 0$, and by using the trivial bound $O(n^2)$ for the derivative,*

$$|P_n(y)| = |P_n(y) - P_n(x)| = O(|x - y|n^2) = O(n^{-B+2}).$$

Delocalization. Assume that $P_n(x) = 0$ and $|a - x| \leq n^{-B} \leq \delta^2$, then $|P_n(a)| = |a - x||P_n'(y)|$ for some y between a and x. On the other hand, $|P_n'(y)| \leq n^2$ for any $y \in [0, 1]$, it follows that

$$|P_n(a)| \leq n^3 \delta^2. \tag{12}$$

To prove Theorem 2.4, we will show that the probability that (10), (11),(12) hold for any fixed point x of I_0 is $O(\delta n^{-C})$. This definitely takes care of (12) and hence (iii) of Theorem 2.4. Since there are $O(\delta^{-1})$ subintervals I, by the union bound we also obtain (i) and (ii) of Theorem 2.4.

In fact we will show the following stronger estimate

Lemma 3.2. *Assume that ξ has type I. Then there is a constant $c > 0$ which depends only on N and ε such that for every $A > 0$ sufficiently large the following holds for every $0 < C_1 \leq C_2$ and $C_1 n^{-A} \leq \delta \leq C_2 n^{-A}$*

$$\sup_{x \in I_0} \mathbf{P}\Big(|P_n(x)| \leq \delta^2\Big) = O(\delta^{1+c}),$$

here the implicit constant may depend on N and c and C_1 and C_2.

(The fact that we have an extra factor of n^3 or n^2 in (10), (11),(12) is not an issue here, since these powers could be included as part of the δ of Lemma 3.2.)

Note that by making c slightly smaller it suffices to prove the Lemma for $C_1 = C_2 = 1$, i.e. $\delta = n^{-A}$, which we will assume in the following. We will justify this key lemma in the next section. In the rest of this section, we apply our argument to handle distributions of Type II.

3.3. **Proof of Theorem 2.6.** By following the same argument, and by (9), it is enough to show the following analog of Lemma 3.2 for Type II variables. Recall that we are working under the assumption that $\xi_0, \dots, \xi_n$ are uniformly Type II and independent, but they are not required to be identically distributed.

Fact 3.4. *There is a constant $c > 0$ which depends only on p such that for every $A > 0$ sufficiently large the following holds for $\delta = n^{-A}$*

$$\sup_{x \in I_0'} \mathbf{P}\Big(|P_n(x)| \leq \delta^2\Big) = O(\delta^{1+c}).$$

Thanks to the analytic properties of Type II variables, this statement is much easier to prove than Lemma 3.2; the details follow.

Proof of Fact 3.4. For any $I \subset \mathbb{R}$, by Hölder's inequality

$$P(\xi_0 \in I) = O(|I|^{1-1/p}) \quad .$$

Thus, by conditioning on $\xi_1, \dots, \xi_n$, for any x we have

$$\mathbf{P}(|\sum_{i=0}^{n} \xi_i x^i| \leq \delta^2) = \mathbf{P}(-\delta^2 - \sum_{i=1}^{n} \xi_i x^i \leq \xi_0 \leq \delta^2 - \sum_{i=1}^{n} \xi_i x^i) = O(\delta^{2(1-1/p)}).$$

The desired conclusion follows immediately if $p > 2$. To handle the general case, let ρ_j denote the density of the distribution of ξ_j, which is p-integrable for $p > 1$ by the given assumption. Since $\int \rho_j(x)dx = 1$, it follows immediately via convexity that ρ_j is also q-integrable for every $q \in [1, p]$ and furthermore

$$\sup_j \|\rho_j\|_q = O_q(1)$$

thanks to the fact that ξ_j's are uniformly Type II. For convenience, let C_q denote the right hand side in the estimate above.

Now, let k be a large integer that depends only on p such that $k/(k-1) < p$. By Young's convolution inequality and an induction over k, it is clear that for any family of functions $g_0, \dots, g_{k-1}$ we have

$$\|g_0 * \cdots * g_{k-1}\|_\infty \leq \prod_{j=0}^{k-1} \|g_j\|_{k/(k-1)}.$$

Consider the random variable $R_k = R_k(x) = \xi_0 + x\xi_1 + \cdots + x^{k-1}\xi_{k-1}$. Since ξ_j's are independent, the density of R_k (which we will denote by r_k) equals to the convolution of the density of ξ_0, $x\xi_1$, $\dots$, $x^{k-1}\xi_{k-1}$. Let g_j denote the density of $x^j\xi_j$, clearly $g_j(t) = x^{-j}\rho_j(t/x^j)$, and

$$\begin{aligned}\|g_j\|_q &= (\int x^{-jq}|\rho_j(t/x^j)|^q dt)^{1/q} \\ &= x^{-j(q-1)/q}\|\rho_j\|_q.\end{aligned}$$

Consequently,

$$\begin{aligned}\|r_k\|_\infty = \|g_0 * \cdots * g_{k-1}\|_\infty &\le \prod_{j=0}^{k-1} x^{-j/k}\|\rho_j\|_{k/(k-1)} \\ &\le (C_{k/(k-1)})^k x^{-(k-1)/2} = O_k(x^{-(k-1)/2}).\end{aligned}$$

Recall that for $x \in I_0'$ we have $\frac{1}{4}n^{-2C_1} \le |x| \le 4n^{2C_1}$. Therefore for every $x \in I_0'$ we have

$$\|r_k\|_\infty = O(n^{C_2}),$$

where C_2 is a finite constant depending only on p and C_1.

Now, for every $n \ge k$ (recall that k is a constant) and $x \in I_0'$ we have

$$\begin{aligned}\mathbf{P}(|P_n(x)| \le \delta^2) &= \mathbf{P}(-\delta^2 - \sum_{i=k}^n \xi_i x^i \le R_k(x) \le \delta^2 - \sum_{i=k}^n \xi_i x^i) \\ &= O(\delta^2\|r_k\|_\infty) = O(n^{C_2}\delta^2).\end{aligned}$$

Thus by choosing A sufficiently large we obtain the desired estimate (with $\delta = n^{-A}$ and any $c < 1$). □

4. Proof of Lemma 3.2: bounds on small value probability for P_n

By making c smaller if necessary, it suffices to prove the Lemma for $\delta = n^{-A}$ where A is sufficiently large. Also, as indicated before, B will be chosen such that $\delta^2 \gg n^{-B}$ (for instance $B = 2A + 10$). Fix $x \in I_0$, all the implicit constants below are independent of x.

We divide I_0 into $(1/(N+1), 1 - \log^2 n/n] \cup (1 - \log^2 n/n, 1 - n^{-2+\varepsilon}]$ and prove the lemma for x inside each interval separately. For the first interval $(1/(N+1), 1 - \log^2 n/n]$, we will present a proof for the Bernoulli case (i.e. $N = 1$) first to demonstrate the main ideas, and then modify the method for uniform distributions later. Our treatment for $(1 - \log^2 n/n, 1 - n^{-2+\varepsilon}]$ works for both settings.

4.1. Proof for $1/(N+1) < x \le 1 - n^{-1}\log^2 n$, the Bernoulli case.

Roughly speaking, the proof exploits the lacunary property of the sequence $\{1, x, \dots, x^n\}$ in this case.

Let $\ell \in \mathbb{Z}$ be such that

$$x^\ell < 1/2 \le x^{\ell-1}.$$

As $x \leq 1 - n^{-1}\log^2 n$, we must have

$$\ell = O(n/\log^2 n) = o(n).$$

Note that if $x < 1/\sqrt{2}$ then $\ell = 2$. As the treatment for this case is a bit more complicated, we postpone it for the moment. In the sequel we assume that $x > 1/\sqrt{2}$, and thus $\ell \geq 3$.

Treatment for $1/\sqrt{2} \leq x \leq 1 - n^{-1}\log^2 n$. Let k be the largest integer such that

$$x^{\ell k} \geq \delta^2 \equiv n^{-2A}$$

In other words,

$$k = \lfloor \frac{(2A)\log n}{\ell \log(1/x)} \rfloor.$$

where $\lfloor x \rfloor$ denote the largest integer that does not exceed x.

As $x \leq 1 - n^{-1}\log^2 n$ and n is sufficiently large, it follows that $k\ell$ is strictly less than n and k is at least $\Omega(\log n)$, one has the following trivial bound

$$k \geq 10.$$

We say that a finite set X of real numbers is γ-separated if the distance between any two elements of X is at least γ.

Claim 4.2. *The set of all possible values of $\sum_{1\leq j\leq k} \varepsilon_j x^{j\ell}, \varepsilon_j \in \{-1, 1\}$ is $2x^{k\ell}$-separated.*

Proof. (of Claim 4.2) Take any two elements of the set. Their distance has the form $2|\varepsilon_{m_1} x^{m_1\ell} + \cdots + \varepsilon_{m_j} x^{m_j\ell}|$ for some $1 \leq m_1 < \cdots < m_j \leq k$. As $x^\ell < 1/2$, this distance is more than $2x^{k\ell}$. □

Using Claim 4.2, we have

$$\sup_R \mathbf{P}_{\xi_{j\ell}, 1\leq j\leq k}(|\sum_{j=1}^{k} \xi_{j\ell} x^{j\ell} + R| \leq x^{k\ell}) \leq 2^{-k}. \tag{13}$$

By conditioning on other coefficients ξ_m's i.e. $m \notin \{\ell, 2\ell, \ldots, k\ell\}$, it follows that

$$\mathbf{P}(|P_n(x)| \leq \delta^2) \leq 2^{-k}. \tag{14}$$

Recall that $x^\ell < 1/2 \leq x^{\ell-1}$. Using the fact that $\ell \geq 3$ and $k \geq 10$, we obtain $(\ell-1)k \geq \frac{3}{5}\ell(k+1)$. It follows that

$$2^{-k} \leq x^{(\ell-1)k} \leq x^{3(k+1)\ell/5} \leq (\delta^2)^{3/5} = \delta^{6/5}.$$

Therefore

$$\mathbf{P}(|P_n(x)| \leq \delta^2) = O(\delta^{6/5})$$

as desired. This completes the treatement of the case $1/\sqrt{2} \leq x \leq 1 - n^{-1}\log^2 n$.

Treatment for $1/2 + c_0 < x < \sqrt{1/2}$. Let c_0 be a small positive constant. We show that the treatment above also carries over for this range of x with a minor modification.

As $x < \sqrt{1/2}$ and $x^\ell < 1/2 \leq x^{\ell-1}$, we must have $\ell = 2$ for all x in this range. Recall that the integer k was chosen so that

$$x^{2(k+1)} < \delta^2 \leq x^{2k} \quad .$$

By following Claim 4.2, we again arrive at (13) and (14).

Now, as $x \geq 1/2 + c_0$, we have $1/2 \leq x^{1+c_1}$ for some small positive constant c_1 depending on c_0. As such, using the fact that k has order $\log n$, we have $\frac{k}{k+1} \geq \frac{1/2+c_1/4}{1/2+c_1/2}$ for n sufficiently large. It follows that

$$\begin{aligned} 2^{-k} \leq x^{(1+c_1)k} &= (x^{2k})^{1/2+c_1/2} \\ &\leq (x^{2(k+1)})^{1/2+c_1/4} < (\delta^2)^{1/2+c_1/4} \quad . \end{aligned}$$

We obtain

$$\mathbf{P}(|P_n(x)| \leq \delta^2) = O(\delta^{1+c_1/2})$$

as desired.

Treatment for $1/2 < x < 1/2+c_0$. Recall that in this case $\ell = 2$. For this range of x we introduce the following improvement of Claim 4.2.

Claim 4.3. *Assume that $1/2 < x < 1/2 + c_0$ and c_0 is sufficiently small. Then the set of all possible values of $\sum_{j=0}^{k} \varepsilon_{2j} x^{2j} + \sum_{j=0}^{\lfloor k/8 \rfloor} \varepsilon_{8j+1} x^{8j+1}, \varepsilon_i \in \{-1, 1\}$, is $x^{2k}/8$-separated.*

Proof. (of Claim 4.3)

The distance between any two terms is at least

$$2x^{2k}[1 - x - \sum_{j=1}^{\infty} x^{2j} - \sum_{j=0}^{\infty} x^{8j+1}] > x^{2k}/8,$$

where we used the fact that the factor within the bracket is at least $1/16$, provided that c_0 is chosen sufficiently small. □

Using Claim 4.3, we obtain the following slight improvement of (13)

$$\sup_R \mathbf{P}_{\xi_{2j}, 0 \leq j \leq k}(|\sum_{j=0}^{k} \xi_{2j} x^{2j} + \sum_{j=0}^{\lfloor k/8 \rfloor} \xi_{8j+1} x^{8j+1} + R| \leq x^{2k}/8) \leq 22^{-k-\lfloor k/8 \rfloor} \leq 2^{-9k/8+2}. \tag{15}$$

Now, as k is order $\log n$, by taking n large we have

$$2^{-9k/8} \leq x^{9k/8} \leq x^{10(k+1)/9} \leq (\delta^2)^{5/9} = \delta^{10/9} \quad ,$$

which implies the desired conclusion.

4.4. **Proof for $1/(N+1) < x \leq 1 - n^{-1}\log^2 n$, the uniform case.** The $N = 1$ case was treated before, so we only consider $N > 1$.

As before, let ℓ be integer such that

$$x^\ell < \frac{1}{2N+1} \leq x^{\ell-1} \quad . \tag{16}$$

Since $x > 1/(N+1)$, it follows that $\ell \geq 2$, and

$$\ell = O_N(\frac{n}{(\log n)^2}).$$

Let k be the largest integer such that

$$x^{\ell k} \geq \delta^2 = n^{-2A}$$

We first show that $k \geq \Omega_N(\log n)$ while $\ell k = o_N(n)$. In deed, by definition we have

$$k = \lfloor \frac{2A \log n}{\ell \log(1/x)} \rfloor.$$

Since $\log(1-a) < -a$ for every $a \in (0,1)$, it follows that

$$\log(1/x) \geq -\log(1 - \frac{(\log n)^2}{n}) \geq \frac{(\log n)^2}{n}$$

therefore

$$k \leq 1 + O(\frac{n}{\ell \log n})$$

Since $\ell = o(n)$, it follows that $k\ell < n$ for n sufficiently large. Furthermore, it follows from (16) that

$$\frac{1}{2N+1} > x^\ell \geq \frac{x}{2N+1} \geq \frac{1}{(N+1)(2N+1)}.$$

Therefore

$$0 < \ell \log(1/x) \leq O_N(1).$$

Hence $k \geq c_N \log n$ for some c_N depending only on N.

Consider the sequence $\sum_{1 \leq j \leq k} \varepsilon_j x^{j\ell}$ where $\varepsilon_j \in \{\pm 1, \pm 2, \ldots, \pm N\}$. We'll show the following separation property:

Claim 4.5. *The set of all possible values of $\sum_{1 \leq j \leq k} \varepsilon_j x^{j\ell}$, where $\varepsilon_j \in \{\pm 1, \pm 2, \ldots, \pm N\}$, is $x^{k\ell}$-separated.*

Proof of Claim 4.5. Take any two terms of the sequence. Consider their difference, which has the form $b_{m_1} x^{m_1 \ell} + \cdots + b_{m_j} x^{m_j \ell}$ for some $1 \leq m_1 < \cdots < m_j \leq k$, and $|b_{m_1}| \geq 1$ and $|b_{m_2}|, \ldots, |b_{m_j}| \leq 2N$. As $x^\ell < 1/(2N+1)$, this difference is more than $x^{k\ell}$. □

It follows that for every R

$$\mathbf{P}(|\sum_{j=1}^{k} \xi_{j\ell} x^{j\ell} + R| \leq x^{kl}) \leq 3(\frac{1}{2N})^k.$$

Now, using the independence of $\xi_0, \ldots, \xi_n$ and by conditioning on ξ_j's with $j \notin \{\ell, 2\ell, \ldots, k\ell\}$, we obtain

$$\mathbf{P}(|P_n(x)| \leq x^{k\ell}) \leq 3(\frac{1}{2N})^k.$$

Thus, by the choice of k, we obtain the key bound

$$\mathbf{P}(|P_n(x)| \leq \delta^2) \leq 3(2N)^{-k}. \tag{17}$$

Next, consider two cases:

Case 1: $1/\sqrt{2N+1} \le x \le 1 - (\log n)^2/n$**.** Since $x^\ell < \frac{1}{2N+1} \le x^{\ell-1}$, it follows that $\ell \ge 3$. Thus,

$$\begin{aligned}(2N)^{-k} &= (1/(2N+1))^{k\log(2N)/\log(2N+1)} \\ &\le x^{(\ell-1)k\log(2N)/\log(2N+1)}.\end{aligned}$$

Let $\gamma_N := \frac{2\log(2N)}{3\log(2N+1)}$. As $\ell - 1 \ge 2\ell/3$, we obtain

$$\begin{aligned}\mathbf{P}(|P_n(x)| \le n^{-2A}) &\le 3(x^{\ell(k+1)})^{\gamma_N k/(k+1)} \\ &\le 3(\delta^2)^{\gamma_N k/(k+1)}.\end{aligned}$$

Since k is controlled below by some $c_N \log n$, we could make $k/(k+1)$ arbitrarily close to 1 by taking n large (independent of α). Thus it suffices to show that

$$\gamma_N > \frac{1}{2}.$$

But it is clear that this holds for every $N \ge 2$. Indeed, consider the function defined on $(0, \infty)$

$$\begin{aligned}f(x) &= 4\log(2x) - 3\log(2x+1) \\ f'(x) &= \frac{4}{x} - \frac{6}{2x+1} = \frac{2x+4}{x(2x+1)} > 0,\end{aligned}$$

so for $x \ge 2$ we have $f(x) \ge f(2) > 0$.

Case 2: $1/(N+1) < x < 1/\sqrt{2N+1}$**.** It follows that $\ell = 2$. Also,

$$\begin{aligned}(2N)^{-k} &= (1/(N+1))^{k\log(2N)/\log(N+1)} \\ &\le x^{k\log(2N)/\log(N+1)}.\end{aligned}$$

Let $\beta_N = \frac{\log(2N)}{2\log(N+1)}$, it follows from (17) that

$$\mathbf{P}(|P_n(x)| \le \delta^2) \le 3(x^{2(k+1)})^{\beta_N k/(k+1)}.$$

By choice of k, $x^{2(k+1)} < \delta^2 \le x^{2k}$, therefore

$$\mathbf{P}(|P_n(x)| \le \delta^2) \le 3(\delta^2)^{\beta_N k/(k+1)}.$$

As before, by choosing n large (independent of x) we could ensure that $k/(k+1)$ is arbitrarily close to 1. Therefore it suffices to show that

$$\beta_N > \frac{1}{2},$$

which is clear for $N > 1$.

4.6. **Roots behaviour in** $1 - \log^2 n/n \le x \le 1 - n^{-2+\varepsilon}$ **for the uniform case.** Our treatment of this interval is more difficult as here the terms x^i are comparable. Our main tool is a following theorem, which is an example of the recently developed machinery of Inverse Littlewood-Offord theorems (see [25] for an introduction).

Theorem 4.7. *Fix positive integers C, C' and $0 < \varepsilon_0 < 1$, and assume that*

$$\rho = \sup_{a \in \mathbb{R}} \mathbf{P}\left(|\sum_i \xi_i v_i - a| \le \beta \right) \ge n^{-C'},$$

for some real numbers $v_1, \dots, v_n$, where ξ_i are iid random variables of uniform distribution with fixed parameter N, and $\beta > 0$ can depend on n. Then for any number n' between n^{ε_0} and n, with n sufficiently large, there exists a proper symmetric generalized arithmetic progression Q, that is $Q = \{\sum_{i=1}^r x_i g_i : x_i \in \mathbf{Z}, |x_i| \le L_i\} \subset \mathbb{R}$, such that

(i) Q is C-proper, i.e. the elements of the set $CQ = \{\sum_{i=1}^r k_i g_i : k_i \in \mathbf{Z}, |k_i| \le CL_i\}$ are all distinct.

(ii) Q has small rank, $1 \le r = O(1)$, and small cardinality

$$|Q| = O(\rho^{-1} \ell_0^{1-r}) = O(\rho^{-1}),$$

where the implied constants here depend on C, C', N and ε_0, and $\ell_0 = \sqrt{n'/\log^2 n}$.

(iii) For all but at most n' elements v of $\{v_1, \dots, v_n\}$, there exists $q \in Q$ such that

$$|q - v| \le T_0 \beta / \ell_0,$$

where $T_0 = \Theta(1)$ independent of C.

(iv) The number $T_0\beta/\ell_0 \in Q$, i.e., there exist $|k_1| \le L_1, \dots, |k_r| \le L_r$ such that

$$T_0\beta/\ell_0 = \sum_i k_i g_i.$$

(v) All steps g_i of Q are integral multiples of $T_0\beta/\ell_0$.

We will provide the deduction of Theorem 4.7 from [25, Theorem 2.9] in Appendix B.

It follows from the C-properness (i) of Q that for any $t \in \mathbf{Z}, 0 < t \le C$, the equation

$$k_1(t)g_1 + \dots + k_r(t)g_r = t(T_0\beta/\ell_0), k_i(t) \in \mathbf{Z}, |k_i(t)| \le CL_i \tag{18}$$

has a unique solution $(k_1(t), \dots, k_d(t)) = t \cdot (k_1, \dots, k_d)$.

Now fix $x \in (1 - \log^2 n/n, 1)$. Choose the largest n_0 so that $x^{n_0} \ge 1/10$, thus

$$n_0 \ge n/\log^2 n.$$

In the sequel, set $\varepsilon_0 := \varepsilon/2$ and

$$n' := n^{\varepsilon_0} \text{ and } \beta := \alpha(1-x)^A \ell_0,$$

where α to be chosen sufficiently small depending on A.

We will prove the following crucial bound.

Lemma 4.8. *We have*

$$\rho = \sup_{r\in\mathbb{R}} \mathbf{P}\Big(|\sum_{i=0}^{n_0} \xi_i x^i - r| \le \beta\Big) = O((n^{1-\varepsilon_0}/\log^2 n)^{-A}).$$

Proof of Lemma 4.8. Without loss of generality assume A is an integer. Assume otherwise that

$$\rho \ge C_1(n^{1-\varepsilon_0}/\log^2 n)^{-A} \tag{19}$$

for some sufficiently large constant C_1 to be chosen depending on all other parameters.

Then by Theorem 4.7, all but n^{ε_0} of the elements are $T_0\beta/\ell_0$-close to a proper GAP Q of rank $r = O(1)$ and size $O(\rho^{-1})$. Furthermore, as noticed, the generators of Q can be chosen to be integral multiples of $T_0\beta/\ell_0$.

Let I be the collection of indices $i \le n_0$ where x^i can be well-approximated by the elements of Q as stated in Theorem 4.7. Then as $|I| \ge n_0 - n^{\varepsilon_0}$, I contains a discrete interval of length $\lfloor n^{1-\varepsilon_0}/\log^2 n\rfloor - 1$, which we denote by $I_0 = \{i_0, \dots, i_0 - \lfloor n^{1-\varepsilon_0}/\log^2 n\rfloor + 2\}$. (Note that the symbol I_0 was used for a different interval previously, which should not be confused with the current setting.)

For any i from I_0 such that $i - A \in I_0$, consider the sequence $x^i, \dots, x^{i-A}$, together with their approximations $q_i, \dots, q_{i-A}$ from Q. By the choice of β,

$$\frac{\alpha^{-1}\beta}{10\ell_0} \le \sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}x^{i-k} = x^{i-A}(1-x)^A \le \frac{\alpha^{-1}\beta}{\ell_0}.$$

As such, by (ii)

$$\frac{\alpha^{-1}\beta}{10\ell_0} - 2^A T_0\beta/\ell_0 \le \sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}q_{i-k} \le \frac{\alpha^{-1}\beta}{\ell_0} + 2^A T_0\beta/\ell_0. \tag{20}$$

With the choice $\alpha = 2^{-A-5}T_0$, one guarantees that $\alpha^{-1} > 10 \times 2^A T_0$, and thus the LHS of (20) is strictly positive. After choosing α, we choose $C = 2^{A+10}T_0$ in Theorem 4.7 so that $C > (\alpha^{-1}+2^A)T_0$, the constant in the RHS of (20).

Next, assume that $q_j = c_{j1}g_1 + \cdots + c_{jd}g_d$ for $|c_{j1}| \le L_1, \dots, |c_{jd}| \le L_d$. Then it follows from the choice of C and from (20) that

$$0 < (\sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}c_{i-k,1})g_1 + \cdots + (\sum_{k=0}^{A}(-1)^{A-k}\binom{A+1}{k}c_{i-k,d})g_d < C\beta/\ell_0.$$

Consequently, recalling that all the generators g_i are integral multiple of $T_0\beta/\ell_0$, there exists $0 < t \le C, t \in \mathbf{Z}$ such that

$$(\sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}c_{i-k,1})g_1 + \cdots + (\sum_{k=0}^{A}(-1)^{A-k}\binom{A+1}{k}c_{i-k,d})g_d = t\beta/\ell_0.$$

It thus follows from (18) that

$$\sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}c_{i-k,1} = tk_1 \wedge \cdots \wedge \sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}c_{i-k,d} = tk_d.$$

In summary, we obtain the following key property for all $i \in I_0$ and $i \geq A$,

$$\sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}c_{i-k,1} \in \{k_1,\ldots,Ck_1\} \wedge \cdots \wedge \sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}c_{i-k,d} \in \{k_d,\ldots,Ck_d\}.$$

As $k_1,\ldots,k_d$ cannot be all zero, without loss of generality, assume that $k_1 > 0$. Thus for every $i \in I_0$ such that $i - A \in I_0$ we have

$$1 \leq k_1 \leq \sum_{k=0}^{A}(-1)^{A-k}\binom{A}{k}c_{i-k,1} \leq Ck_1. \tag{21}$$

We next require the following observation.

Claim 4.9. *Assume that $\{x_i\}_{i=0}^m$ is a sequence of real numbers which satisfy the following inequality for all $A \leq i \leq m$*

$$1 \leq \sum_{k=0}^{A}(-1)^{k}\binom{A}{k}x_{i-k}.$$

Then there exist $0 \leq i,j \leq m$ such that

$$|x_i - x_j| \geq C_A m^A,$$

where $C_A > 0$ depends on A.

Proof. (Proof of Claim 4.9) Define

$$\Delta^0(x_i) := x_i \text{ and } \Delta^k(x_i) := \Delta^{k-1}(x_{i-1}) - \Delta^{k-1}(x_i).$$

By the assumption and Pascal's triangle identity

$$1 \leq \sum_{j=0}^{A}(-1)^{j}\binom{A}{j}x_{i-j} = \Delta^A(x_i), \forall A \leq i \leq m.$$

It follows that $\Delta^{A-1}(x_{i-1}) \geq \Delta^{A-1}(x_i) + 1$. Thus, there are at least $(m-A)/4$ consecutive indices $i \geq A$ such that the corresponding $\Delta^{k-1}(x_i)$ have the same signs and absolute value at least $(m-A)/4$. Without loss of generality, we can assume that all of them are at least $(m-A)/4$. Repeat the argument with $A-1$ and the above subsequence. After A repetitions, we obtain a sub interval I of length $(m-A)/4^A$ of $[A,m]$, where all $x_i, i \in I$ have the same signs and absolute value at least $((m-A)/4)^A$. □

Applying Claim 4.9 with $m = \lfloor n^{1-\varepsilon_0}/\log^2 n \rfloor - 2$ and $\{x_i\}_{i=0}^m := \{c_{i_0-i,1}\}_{i=0}^m$, we obtain $L_1 \geq C_A m^A$, and so

$$|Q| \geq C_A(n^{1-\varepsilon_0}/\log^2 n)^A.$$

This contradicts with the bound $|Q| = O(\rho^{-1})$ from Theorem 4.7 and with the bound $\rho \geq C_1(n^{1-\varepsilon_0}/\log^2 n)^{-A}$ from (19) because C_1 is sufficiently large. This completes the proof of Lemma 4.8. □

Now we conclude the subsection by proving Theorem 2.4 for the interval $(1 - \log^2 n/n, n^{-2+\varepsilon}]$.

Proof of Lemma 3.2 for $1 - \log^2 n/n < x \leq 1 - n^{-2+\varepsilon}$. We need to show the existence of $c > 0$ (that depends only on ε and N) such that for every $A > 0$ sufficiently large

$$P(|P_n(x)| \leq \delta^2) = O(\delta^{1+c}).$$

We will apply Lemma 4.8 with $\varepsilon_0 = \varepsilon/2$, and let α be the corresponding constant. By making c smaller if necessary, it suffices to prove Lemma 3.2 for

$$\delta = \sqrt{\alpha} n^{-A(1-\varepsilon/2)+\varepsilon_0/8}$$

where A is sufficiently large (instead of requiring $\delta = n^{-A}$ as before.)

As $x \in (1 - \log^2 n/n, 1 - n^{-2+\varepsilon}]$, one can verify that

$$\alpha(1-x)^A \ell_0 \quad = \quad \alpha(1-x)^A \sqrt{n'}/\log n \quad \geq \quad \alpha n^{-A(2-\varepsilon)+\varepsilon/8} = \delta^2.$$

Thus, by Lemma 4.8,

$$\begin{aligned} \mathbf{P}\Big(|P(x)| \leq \delta^2\Big) \quad \leq \quad \sup_{r\in\mathbb{R}} \mathbf{P}\Big(|\sum_{i=0}^{n_0} \xi_i x^i - r| \leq \alpha(1-x)^A \ell_0\Big) &= O\big((n^{1-\varepsilon_0}/\log^2 n)^{-A}\big) \\ &= O(n^{-A(1-3\varepsilon/8)}) \\ &= O(\delta^{1+c}) \quad , \end{aligned}$$

for sufficiently small c, provided that A is sufficiently large. □

5. Proof of Theorem 2.3: no double root at the edge

Set $t := 1 - x$. Then $0 \leq t \leq n^{-2+\varepsilon}$. For every $i = 0, 1, \ldots, n$, write

$$x^i = (1-t)^i = \sum_{0\leq k\leq n} (-1)^k \binom{i}{k} t^k,$$

where $\binom{i}{k} = 0$ if $i < k$. Consequently

$$P_n(x) = \sum_{k=0}^{n} (-1)^k \Big(\sum_i \binom{i}{k} \xi_i\Big) t^k$$

$$P_n'(x) = \sum_{k=1}^{n} (-1)^{k-1} \Big(\sum_i i\binom{i-1}{k-1} \xi_i\Big) t^{k-1} = \sum_{k=1}^{n} (-1)^{k-1} k \Big(\sum_i i\binom{i}{k} \xi_i\Big) t^{k-1}$$

Notice easily that, with probability $1 - \exp(-\Omega(\log^2 n))$,

$$|\sum_i \binom{i}{k} \xi_i| = n^{k+1/2} \log^{O(1)} n, \forall 1 \leq k \leq n. \tag{22}$$

Let $\mathcal{E}$ denote this event, on which we will condition for the rest of our argument. Thus

$$\sum_{k\geq k_0}(-1)^k(\sum_{0\leq i\leq n}\binom{i}{k}\xi_i)t^k = (nt)^{k_0}n^{1/2}\log^{O(1)} n. \tag{23}$$

In particular, with $k_0 = 1$,

$$\sum_{k\geq 1}(-1)^k(\sum_{0\leq i\leq n}\binom{i}{k}\xi_i)t^k = O(n^{-1/2+\varepsilon}\log^{O(1)} n),$$

therefore $\sum_i \xi_i = P_n(x) + o(1)$. But as ξ_i takes integer values and $|P_n(x)| = o(1)$, therefore by taking n large we must have

$$\sum_{i=0}^{n}\xi_i = 0. \tag{24}$$

After replacing (24) into $|P_n(x)|$, we obtain

$$t|\sum_{k\geq 1}(-1)^k(\sum_{0\leq i\leq n}\binom{i}{k}\xi_i)t^{k-1}| = |P_n(x)| = O(n^{-B}). \tag{25}$$

Now we consider the assumption that $|P_n'(x)| \leq n^{-B}$, from which we infer that

$$|P_n'(x)| = |\sum_{k\geq 1}(-1)^k k(\sum_{0\leq i\leq n}\binom{i}{k}\xi_i)t^{k-1}| = \tag{26}$$

$$= |-\sum_{0\leq i\leq n} i\xi_i + \sum_{k\geq 2}(-1)^k k(\sum_{0\leq i\leq n}\binom{i}{k}\xi_i)t^{k-1}| = O(n^{-B}).$$

We next consider two cases.

Case 1. $t \leq n^{-8}$. It follows easily from (22) and (26) that

$$\sum_{0\leq i\leq n} i\xi_i = o(1) \quad .$$

Case 2. $t \geq n^{-8}$. As $B \geq 16$, it follows from (25) that

$$|\sum_{k\geq 1}(-1)^k(\sum_{0\leq i\leq n}\binom{i}{k}\xi_i)t^{k-1}| = |-\sum_{0\leq i\leq n} i\xi_i + \sum_{k\geq 2}(-1)^k(\sum_{0\leq i\leq n}\binom{i}{k}\xi_i)t^{k-1}| \leq n^{-B/2}. \tag{27}$$

Combining (26) and (28) together to eliminate the term corresponding to $k = 2$, and also by (22)

$$\sum_{0\leq i\leq n}\binom{i}{1}\xi_i + O(n^{7/2}\log^{O(1)} n)t^2 = O(n^{-B/2}). \tag{28}$$

Thus, using the fact that $t \leq n^{-2+\epsilon} \leq n^{-15/8}$ with $\varepsilon \leq 1/8$, it follows that

$$|\sum_i i\xi_i| = o(1).$$

As ξ_i takes integer values, it follows from both cases that

$$\sum_i i\xi_i = 0. \tag{29}$$

6. Application: proof of Theorem 1.5

Since ξ has no atom at 0, $P_n(0) \neq 0$ with probability 1. First of all, the contribution towards the expectation at the points ± 1 are negligible owing to the following elementary estimates.

Claim 6.1. *We have*

$$\mathbf{P}\Big(P_n(1)P_n(-1) = 0\Big) = O(1/\sqrt{n})$$
$$\mathbf{P}\Big(P_n(1) = 0 \wedge P_n'(1) = 0\Big) = O(1/n^2)$$
$$\mathbf{P}\Big(P_n(-1) = 0 \wedge P_n'(-1) = 0\Big) = O(1/n^2).$$

Note that the above claim is trivial in the Type II setting, so it remains to show these estimates for Type I. The first estimate clearly follows from the classical Erdős-Littlewood-Offord bound. We refer the reader to Claim A.5 for a short proof of the remaining two estimates.

Using Claim 6.1, it follows that

$$\begin{aligned}\mathbf{E}N_{n,\xi}\{-1,1\} &= O(1/\sqrt{n}) + nO(1/n^2)\\ &= O(1/\sqrt{n}) \quad .\end{aligned}$$

Now, using the fact that the (real) zero sets of $P_n(x)$ and $x^n P_n(1/x)$ have the same distribution, it follows that

$$\begin{aligned}\mathbf{E}N_{n,\xi} &= 2\mathbf{E}N_{n,\xi}(-1,1) + \mathbf{E}N_{n,\xi}\{-1,1\}\\ &= 2\mathbf{E}N_{n,\xi}(-1,1) + O(1/\sqrt{n}) \quad .\end{aligned}$$

In the following, we will consider the number of real roots in $(0,1)$, and show that

$$\mathbf{E}N_{n,\xi}(0,1) = \frac{1}{2\pi}\log n + C + o(1) \tag{30}$$

for some $C = C(\xi)$.

To estimate $\mathbf{E}N_{n,\xi}(-1,0)$, we consider the random polynomial

$$\widetilde{P}_n(x) := P_n(-x) = \xi_0 - \xi_1 x + \xi_2 x^2 + \cdots + (-1)^n \xi_n x^n$$

and let $\widetilde{N}_{n,\xi}$ denote its number of real zeros. By definition, we have

$$\mathbf{E}N_{n,\xi}(-1,0) = \mathbf{E}\widetilde{N}_{n,\xi}(0,1)$$

so we need to estimate the average number of real zeros for $\widetilde{P}_n$ in $(0,1)$.

Now, Type I distributions are symmetric so in that setting the zero sets of $\widetilde{P}_n$ and P_n would have the same distribution, therefore $\mathbf{E}\widetilde{N}_{n,\xi} = \mathbf{E}N_{n,\xi}$ and one obtains the same estimate as above for $\mathbf{E}N_{n,\xi}(-1,0)$.

For Type II distributions, the argument we use below for estimating $\mathbf{E}N_{n,\xi}(0,1)$ could be applied to estimate $\mathbf{E}\widetilde{N}_{n,\xi}(0,1)$. Most importantly, our result on non-existence of double roots (Theorem 2.6) does not require the coefficients $\xi_0,\dots,\xi_n$ to have identical distributions, and could be applied to $\widetilde{P}_n$. Also the cited ingredient that we used below (Theorem 6.3) does not require $\xi_0,\dots,\xi_n$ to have identical distributions. With cosmetic changes, one could use the argument below to $\widetilde{P}_n$ and obtains a similar estimate for $\mathbf{E}\widetilde{N}_{n,\xi}(0,1)$, which is the same as $\mathbf{E}N_{n,\xi}(-1,0)$, and conclude the proof of Theorem 1.5. For the rest of the section we will be focusing on (30).

6.2. **Comparison lemmas.** Our first tool is the following corollary from [31, Theorem 5.8].

Theorem 6.3. *There is a positive constant α such that the following holds. Let $\varepsilon > 0$ be an arbitrary small constant and $\xi_0,\dots,\xi_n$ be independent random variables with mean 0, variance 1 and uniformly bounded $(2+\varepsilon)$-moment. There is a constant $C_1 = C_1(\varepsilon)$ such that for any $n \geq C_1$ and any interval $I := (1-r,a) \subset (1-n^{-\varepsilon},1]$, with $a \leq 1$*

$$\mathbf{E}N_{n,\xi_0,\dots,\xi_n}I = \mathbf{E}N_{n,N(0,1)}I + O(r^\alpha), \tag{31}$$

where the implicit constant in $O(.)$ depends only on α and ε.

Next, for convenience, we truncate the random variables $\xi_0,\dots,\xi_n$. Let $d>0$ be a parameter and let $\mathcal{B}_d$ be the event $|\xi_0| < n^d \wedge \dots \wedge |\xi_n| < n^d$. As ξ has mean zero and bounded $(2+\varepsilon_0)$-moment, we have the following elementary bound for $d \geq 1$

$$\mathbf{P}(\mathcal{B}_{2d}^c) = O(n^{1-3d}).$$

In what follows we will condition on $\mathcal{B}_4$. Consider $P_n(x) = \sum_{i=0}^n \xi_i x^i$ and for $m<n$, we set

$$g_m := P_n - P_m = \sum_{i=m+1}^n \xi_i x^i.$$

For any $0 < x \leq 1-r$, a generous Chernoff's bound yields that for any $\lambda > 0$

$$\begin{aligned}
&\mathbf{P}\Big(|g_m(x)| \geq (\lambda+1)n^5 \sqrt{\sum_{i=m+1}^n (1-r)^{2i}}\Big|\mathcal{B}_4\Big) \quad \leq \\
\leq\ &\mathbf{P}\Big(|g_m(x)| \geq (\lambda+1)n^5 \sqrt{\sum_{i=m}^n x^{2i}}\Big|\mathcal{B}_4\Big) \quad \leq \\
\leq\ &2\exp(-\lambda^2/2)\ .
\end{aligned}$$

Since

$$\sum_{i=m+1}^n (1-r)^{2i} \leq (1-r)^{2m+2}\frac{1}{1-(1-r)^2} := s(r,m),$$

it follows that

$$\mathbf{P}(|g_m| \geq (\lambda+1)n^4\sqrt{s(r,m)}|\mathcal{B}_4) \leq 2\exp(-\lambda^2/2). \tag{32}$$

We next compare the roots of P_n and P_m in the interval $(0,1-r)$.

Lemma 6.4 (Roots comparison for truncated polynomials)**.** *Let $r \in (1/n, 1)$ and $m \leq n$ such that $m \geq 4Br^{-1}\log n$, where $B = B(3)$ of Theorem 2.4. Then for any subinterval J of $(0, 1-r]$ one has*

$$|\mathbf{E}N_{n,\xi}J - \mathbf{E}N_{m,\xi}J| = O(m^{-2}) \quad , \tag{33}$$

and the implicit constant depends on N and B only.

To prove Lemma 6.4, we need the following elementary lemma from [26].

Lemma 6.5. *Assume that $F(x) \in C^2(\mathbb{R})$ and $G(x)$ are continuous functions satisfying the following properties*

- *$F(x_0) = 0$ and $|F'(x_0)| \geq \epsilon_1$;*
- *$|F''(x)| \leq M$ for all $x \in I := [x_0 - \epsilon_1 M^{-1}, x_0 + \epsilon_1 M^{-1}]$;*
- *$\sup_{x\in I} |F(x) - G(x)| \leq \frac{1}{4}\epsilon_1^2 M^{-1}$.*

Then G has a root in I.

Proof of Lemma 6.4. Conditioned on $\mathcal{B}_4$, with probability at least $1-2\exp(-\log^2 n/2) \geq 1-n^{-\omega(1)}$ the following holds

$$|P_n(x) - P_m(x)| \leq n^5(\log n + 1)\sqrt{s(r,m)} = n^5(\log n + 1)(1-r)^{m+1}\frac{1}{\sqrt{1-(1-r)^2}} \leq n^{-3B}$$

for all $0 \leq x \leq 1-r$ and n sufficiently large.

By Theorem 2.4 or Theorem 2.6 (with $C = 3$), $|P_n'(x)| \geq n^{-B}$ for all $x \in J$ with probability $1 - O(n^{-3})$. Note that the conditioning on $\mathcal{B}_4$, which holds with probability at least $1 - O(n^{-5})$, will not affect the $O(n^{-3})$ term in this estimate. Applying Lemma 6.5 with $\epsilon_1 = n^{-B}, M = n^3$, $F = P_n, G = P_m$, we conclude that with probability $1 - O(n^{-3})$, for any root x_0 of $P_n(x)$ in the interval $(0, 1-r)$ (which is a subset of $(0, 1-1/n)$), there is a root y_0 of $P_m(x)$ such that $|x_0 - y_0| \leq \epsilon_1 M^{-1} = n^{-B-3}$.

On the other hand, applying (2) of Theorem 2.4 or Theorem 2.6 with $C = 3$, again with probability $1 - O(n^{-3})$ there is no pair of roots of P_n in J with distance less than n^{-B}. It follows that for different roots x_0 we can choose different roots y_0. Furthermore, by (3) of Theorem 2.4 or Theorem 2.6, with probability $1 - O(n^{-3})$, all roots of $P_n(x)$ must be of distance at least n^{-B} from the two ends of the interval. If this holds, then all y_0 must also be inside the interval. This implies that with probability at least $1 - O(n^{-3})$, the number of roots of P_m in J is at least that of P_n. Putting together, we obtain

$$\mathbf{E}N_{m,\xi}J \geq \mathbf{E}N_{n,\xi}J - O(n^{-3})n \geq \mathbf{E}N_{n,\xi}J - O(n^{-2}), \tag{34}$$

where the factor n comes from the fact that P_n has at most n real roots.

Switching the roles of P_n and P_m, noting that as $r \geq 4B\log n/m > 1/m$,

$$J \subset (0, 1-r] \subset (0, 1-1/m).$$

As such, Theorem 2.4 and Theorem 2.6 are also applicable to $P_m(x)$. Argue similarly as above, we also have

$$\mathbf{E}N_{n,\xi}J \geq \mathbf{E}N_{m,\xi}J - (O(m^{-3}) + n^{-3B})m \geq \mathbf{E}N_{m,\xi}J - O(m^{-2}). \tag{35}$$

It follows that

$$|\mathbf{E}N_{n,\xi}J - \mathbf{E}N_{m,\xi}J| = O(m^{-2}).$$

□

6.6. **Control of the error term.** By iterating Lemma 6.4, one can achieve the following.

Corollary 6.7. *Let C_0 be sufficiently large, and let $I = (0, 1 - C_0^{-1})$. Then for any sufficiently large integer L (depending on $B = B(3)$ and C_0) and $n \geq L$*

$$\left|\mathbf{E}N_{n,\xi}I - \mathbf{E}N_{L,\xi}I\right| = O(L^{-1}) \quad ,$$

where the implicit constant depends only on $B(3)$ and the parameter N of ξ.

Proof of Corollary 6.7. Let $r = 1/C_0$. We assume $L > C_0$ so that $r \in (1/n, 1)$ for every $n \geq L$. Define the sequence $\{n_i\}$ with $n_0 = n$ and $n_{i+1} = 1 + \lfloor 4Br^{-1} \log n_i \rfloor$. By ensuring that C_0 is sufficiently large we will have $n_i > n_{i+1}$ unless $n_i = 1$. Thus, the sequence $\{n_i\}_0^k$ is decreasing, and let k be the first index where $n_{k+1} \leq L$. Then

$$\begin{aligned}|\mathbf{E}N_{n,\xi}I - \mathbf{E}N_{L,\xi}I| &\leq \sum_{i=0}^{k-1} |\mathbf{E}N_{n_i,\xi}I - \mathbf{E}N_{n_{i+1},\xi}I| + |\mathbf{E}N_{n_k,\xi}I - \mathbf{E}N_{L,\xi}I| \\ &= O(\sum_{i=0}^{k-1} n_{i+1}^{-2} + L^{-2}) = O(L^{-1}).\end{aligned}$$

Here the last estimate also holds by applying Lemma 6.4 for $n = n_k$ and $m = L$, and clearly $L \geq n_{k+1} > 4Br^{-1} \log n_k$, and $r \in (1/L, 1) \subset (1/n_k, 1)$. □

Next, for each L denote $C_L := \mathbf{E}N_{L,\xi}I = \mathbf{E}N_{L,\xi}[0, 1 - C_0^{-1})$. By Corollary 6.7, $\{C_L\}$ is a Cauchy sequence, thus it has a finite limit.

Corollary 6.8. *For any $C_0 > 0$, there exists $C^* = C^*(C_0) < \infty$ such that*

$$\lim_{n\to\infty} \mathbf{E}N_{n,\xi}(0, 1 - C_0^{-1}) = C^*.$$

6.9. **Control of the main term.** We next turn to the main term by utilizing Theorem 6.3 and Lemma 6.4. Recall the constant α from Theorem 6.3. Let $0 < \varepsilon < \alpha/2$ be a small constant (so that in particular $\varepsilon < 1$), and let $C_1 = C_1(\varepsilon)$ to be the constant of Theorem 6.3.

Lemma 6.10. *Assume that $C_0 > C_1^{\varepsilon}$, then*

$$|\mathbf{E}N_{n,\xi}(1 - {C_0}^{-1}, 1) - \mathbf{E}N_{n,N(0,1)}(1 - {C_0}^{-1}, 1)| = O({C_0}^{-\alpha}),$$

here the implicit constant depends only on N, α, and ε.

We will show the following equivalent statement: assume that $C_0 > C_1$, then

$$|\mathbf{E}N_{n,\xi}(1 - {C_0}^{-\varepsilon}, 1) - \mathbf{E}N_{n,N(0,1)}(1 - {C_0}^{-\varepsilon}, 1)| = O({C_0}^{-\varepsilon\alpha}). \tag{36}$$

Proof of Lemma 6.10. We will justify (36) following [26]. Set $n_0 := n, r_0 = n^{-\varepsilon}$ and define recursively

$$n_i := 1 + \lfloor 4Br_{i-1}^{-1} \log n_{i-1} \rfloor, \text{ and } r_i := n_i^{-\varepsilon}, i \geq 1.$$

It is clear that $\{n_i\}$ and $\{r_i\}$ are, respectively, strictly decreasing and increasing sequences . Let L be the largest index such that $n_L > C_0$. It follows that $C_0 \geq n_{L+1} > 4Bn_L^\varepsilon \log n_L \geq n_L^\varepsilon$, therefore $n_L < C_0^{1/\epsilon}$. It also follows that

$$4B \log n_L < C_0^{1-\epsilon}$$

Redefine $n_{L+1} := 1 + \lfloor C_0 \rfloor$ and $r_{L+1} = C_0^{-\varepsilon}$. It is clear that we still have $n_L \geq n_{L+1} \geq 4Br_L^{-1} \log n_L$.

For $1 \leq i \leq L+1$, define $I_i := (1 - r_i, 1 - r_{i-1}]$. For every $1 \leq j \leq i$ we have $I_i \subset (0, 1 - r_{j-1}]$ while $r_{j-1} \in (1/n_{j-1}, 1)$. Thus, by (33), for any $1 \leq j \leq i$ we have

$$|\mathbf{E}N_{n_{j-1},\xi} I_i - \mathbf{E}N_{n_j,\xi} I_i| = O(n_j^{-2}).$$

By the triangle inequality,

$$|\mathbf{E}N_{n_0,\xi} I_i - \mathbf{E}N_{n_i,\xi} I_i| = O(\sum_{j=1}^{i} n_j^{-2}) = O(n_i^{-1}). \tag{37}$$

Similarly, as standard Gaussian distribution is of type II,

$$|\mathbf{E}N_{n_0,N(0,1)} I_i - \mathbf{E}N_{n_i,N(0,1)} I_i| = O(\sum_{j=1}^{i} n_j^{-2}) = O(n_i^{-1}). \tag{38}$$

On the other hand, note that $n_i \geq C_1$ for $i \leq L+1$, and $I_i = (1 - n_i^{-\varepsilon}, 1 - n_{i-1}^{-\varepsilon}] \subset (1 - n_i^{-\varepsilon}, 1)$. By Theorem 6.3

$$|\mathbf{E}N_{n_i,\xi} I_i - \mathbf{E}N_{n_i,N(0,1)} I_i| \leq O(n_i^{-\epsilon\alpha}). \tag{39}$$

Combining (37) and (39), one obtains

$$|\mathbf{E}N_{n_0,\xi} I_i - \mathbf{E}N_{n_0,N(0,1)} I_i| = O(n_i^{-1} + n_i^{-\epsilon\alpha}). \tag{40}$$

Let $I = \cup_{i=0}^{L+1} I_i$, again by the triangle inequality

$$|\mathbf{E}N_{n,\xi} I - \mathbf{E}N_{n,N(0,1)} I| = O(\sum_{i=0}^{L+1} n_i^{-1} + \sum_{i=0}^{L+1} n_i^{-\epsilon\alpha}).$$

The right end point of I is $1 - n^{-\varepsilon}$, and the left end point is $1 - r_{L+1} = 1 - C_0^{-\varepsilon}$. Furthermore, by definition of the n_i, it is easy to see that $(n_i)_{i=0}^{L+1}$ is lacunary, therefore

$$\sum_{i=0}^{L+1} n_i^{-\epsilon\alpha} = O(n_{L+1}^{-\epsilon\alpha}) = O({C_0}^{-\varepsilon\alpha}) \quad .$$

Thus,

$$|\mathbf{E}N_{n,\xi}(1 - C_0^{-\varepsilon}, 1 - n^{-\varepsilon}]) - \mathbf{E}N_{n,N(0,1)}(1 - C_0^{-\varepsilon}, 1 - n^{-\varepsilon}]| = |\mathbf{E}N_{n,\xi} I - \mathbf{E}N_{n,N(0,1)} I| = O({C_0}^{-\varepsilon\alpha}).$$

Combined with Theorem 6.3,

$$|\mathbf{E}N_{n,\xi}(1 - C_0^{-1}, 1) - \mathbf{E}N_{n,N(0,1)}(1 - C_0^{-1}, 1)| = O({C_0}^{-\varepsilon\alpha}) + O(n^{-\epsilon\alpha}) = O({C_0}^{-\varepsilon\alpha}),$$

proving (36).

□

6.11. **Completing the proof of Theorem 1.5.** It suffices to show that

$$\mathbf{E}N_{n,\xi}(0,1) = \frac{1}{2\pi}\log n + C_\xi + o(1). \tag{41}$$

To this end, we first complement the result of Lemma 6.10 by giving an estimate for $\mathbf{E}N_{n,N(0,1)}(1-C_0^{-1},1)$.

Claim 6.12. *For every $C_0 \in (1,\infty)$ there exists a finite number $B^* = B^*(C_0)$ that depends only on C_0 such that*

$$\mathbf{E}N_{n,N(0,1)}(1-C_0^{-1},1) = \frac{1}{2\pi}\log n + B^* + o_{C_0}(1) \quad ,$$

here the $o_{C_0}(1)$ term is with respect to the limit $n\to\infty$, and the implied constant depends on C_0.

Proof of Claim 6.12. We will use the following formula from [5], which asserts that the following equality holds for every interval $I \subset [0,1]$:

$$\mathbf{E}N_{n,N(0,1)}I = \int_I \frac{1}{\pi}\sqrt{\frac{1}{(t^2-1)^2} - \frac{(n+1)^2t^{2n}}{(t^{2n+2}-1)^2}}dt.$$

In particular, for every fixed $1 < C_0 < \infty$ we have

$$\lim_{n\to\infty}\mathbf{E}N_{n,N(0,1)}[0,1-C_0^{-1}) = \int_0^{1-C_0^{-1}} \frac{1}{\pi(1-t^2)}dt$$

thus we can define

$$B^*(C_0) = \frac{1}{4}C_{Gau} - \int_0^{1-C_0^{-1}} \frac{1}{\pi(1-t^2)}dt$$

here recall that C_{Gau} is the constant in the asymptotics expansion (6) of the Gaussian Kac polynomial. □

It then follows from Lemma 6.10 that

$$\mathbf{E}N_{n,\xi}(1-C_0^{-1},1) = \frac{1}{2\pi}\log n + B^*(C_0) + O(C_0^{-\alpha}) + o_{C_0}(1) \quad ,$$

and in the $O(C_0^{-\alpha})$ term the implicit constant may depend on α,ϵ and ξ.

Combining with Corollary 6.8, we obtain

$$\mathbf{E}N_{n,\xi}(0,1) = \frac{1}{2\pi}\log n + B^*(C_0) + C^*(C_0) + O(C_0^{-\alpha}) + o_{C_0}(1),$$

where $C^* = C^*(C_0)$ is a number depending on C_0.

Replacing C_0 by C_0' and subtract,

$$|B^*(C_0') + C^*(C_0') - B^*(C_0) - C^*(C_0)| = O(C_0^{-\alpha} + C_0'^{-\alpha}) + o_{C_0,C_0'}(1),$$

and sending $n\to\infty$ we obtain

$$|B^*(C_0') + C^*(C_0') - B^*(C_0) - C^*(C_0)| = O(C_0^{-\alpha} + C_0'^{-\alpha})$$

which shows that the function $B^*(C_0) + C^*(C_0)$ tends to a limit as C_0 tends to infinity. Denote this limit by C_ξ, it follows that

$$\mathbf{E}N_{n,\xi}(0,1) = \frac{1}{2\pi}\log n + C_\xi + o(1),$$

concluding the proof.

Remark 6.13. *Note that in this application section, we have applied our roots repulsion results (Theorem 2.4 and Theorem 2.6) only for the interval* $(0, 1 - \log^2 n/n]$.

Acknowledgments. The authors are grateful to the anonymous referees for many valuable suggestions and corrections.

Appendix A. Sharpness of Fact 1.7 and proof of Claim 6.1

We first address Fact 1.7. It is clear that when n is even then $P_n(1), P_n(-1)$ are odd numbers, and hence the polynomial cannot have double roots at these points. We next show the same thing for the case $n = 4k + 1$.

Fact A.1. *Assume that* $n = 4k + 1$*, then* $P_n(x)$ *cannot have double root at* -1 *or* 1.

Proof. (of Fact A.1) Assume that $P_n(1) = P_n'(1) = 0$, then one has

$$\xi_0 + \xi_1 + \cdots + \xi_{4k} + \xi_{4k+1} = 0 \text{ and } \xi_1 + 2\xi_2 + \cdots + 4k\xi_{4k} + (4k+1)\xi_{4k+1} = 0.$$

Consequently,

$$\xi_2 + 2\xi_3 + \cdots + (4k-1)\xi_{4k} + 4k\xi_{4k+1} = \xi_0.$$

However, this is impossible as the RHS is an odd number, while the LHS is clearly an even number for any choice of $\xi_2, \dots, \xi_{4k+1} \in \pm 1$.

The non-existence of double root at -1 can be argued similarly (or just by setting $Q_n(x) := P_n(-x)$.) □

We now give a brief explanation that the probability bound of Fact 1.7 is sharp, up to a multiplicative constant.

Lemma A.2. *Let* ξ *be a Bernoulli random variable and* $n + 1$ *be divisible by 4. Then*

$$\mathbf{P}(P_n(1) = P_n'(1) = 0) = \Omega(n^{-2}).$$

We use the circle method. Let $v_i := (1, i)$, and $V = \{v_0, \dots, v_n\}$. Let $p \gg n$ be a sufficiently large prime. We first write

$$\mathbf{P}(\sum_i \xi_i v_i = 0) = \mathbf{E}_{\xi_0,\dots,\xi_n}\mathbf{E}_{\mathbf{x}\in(\mathbf{Z}/p\mathbf{Z})^2} e_p(\langle \sum_{i=1}^n \xi_i v_i, \mathbf{x}\rangle) = \mathbf{E}_{\mathbf{x}\in(\mathbf{Z}/p\mathbf{Z})^2} \prod_{i=0}^n \cos(2\pi\langle v_i, \mathbf{x}\rangle/p). \tag{42}$$

Observe that $|\cos(\pi 2wx/p)| \le 3/4 + \cos(4\pi wx/p)/4 \le \exp(-\|2wx/p\|^2)$, where $\|.\|$ is the distance to the nearest integer. We are going to analyze the sum

$$\sum_{v_i \in V} \|2\langle \mathbf{x}, v_i\rangle/p\|^2 = \sum_{i=0}^{n} \|(2x_1 + 2ix_2)/p\|^2.$$

Basically, if this sum is quite large, then its distribution in (42) is negligible. We state the following elementary claims whose proofs are left to the reader as an exercise.

Claim A.3. *The following holds.*

- *Assume that* $\|\frac{2x_2}{p}\| \ge \frac{\log^2 n}{n^{3/2}}$, *then*
$$\sum_{i=0}^{n} \|(2x_1 + 2ix_2)/p\|^2 \ge 4\log n.$$
- *Assume that* $\|\frac{2x_2}{p}\| \le \frac{\log^2 n}{n^{3/2}}$ *and* $\|\frac{2x_1}{p}\| \ge \frac{\log^2 n}{n^{1/2}}$, *then*
$$\sum_{i=0}^{n} \|(2x_1 + 2ix_2)/p\|^2 \ge 4\log n.$$

It follows from Claim A.3 that the main term of the sum in (42) is governed by $\|2x_1/p\| \le \log^2 n/n^{1/2}$ and $\|2x_2/p\| \le \log^2 n/n^{3/2}$. Thus either $\|x_1/p\| \le \log^2 n/2n^{1/2}$ or $\|x_1/p + 1/2\| \le \log^2 n/2n^{1/2}$ and $\|x_2/p\| \le \log^2 n/2n^{3/2}$ or $\|x_2/p + 1/2\| \le \log^2 n/2n^{3/2}$. As $4|n+1$, the interested reader is invited to check that the contribution in (42) of one of these four cases are the same. It is thus enough to work with the case $\|x_1/p\| \le \log^2 n/2n^{1/2}$ and $\|x_2/p\| \le \log^2 n/2n^{3/2}$. We are going to show the following.

Lemma A.4.

$$\begin{aligned} S :&= \frac{1}{p^2} \sum_{\|\frac{x_1}{p}\| \le \frac{\log^2 n}{n^{1/2}}, \|\frac{x_2}{p}\| \le \frac{\log^2 n}{n^{3/2}}} \prod_{i=0}^{n} \cos \frac{2\pi(x_1 + ix_2)}{p} \\ &= \frac{1}{p^2} \sum_{\|\frac{x_1}{p}\| \le \frac{\log^2 n}{n^{1/2}}, \|\frac{x_2}{p}\| \le \frac{\log^2 n}{n^{3/2}}} \prod_{i=0}^{n} \Big|\cos \frac{2\pi(x_1 + ix_2)}{p}\Big| \\ &= \Omega(n^{-2}). \end{aligned}$$

The method below gives an exact estimate on the asymptotic constant, but we will not need this fact.

Proof of Lemma A.4. The first equality is trivial, as all cosines are positive in this range of $\mathbf{x}$. Viewing x_1 and x_2 as integers with absolute value at most $p\log^2 n/n^{1/2}$ and $p\log^2 n/n^{3/2}$ respectively. We have

$$\cos \frac{2\pi(x_1 + ix_2)}{p} = 1 - \Big(\frac{1}{2} + o(1)\Big)\frac{4\pi^2(x_1 + ix_2)^2}{p^2} = \exp\Big(-(1/2 + o(1))\frac{4\pi^2(x_1 + ix_2)^2}{p^2}\Big).$$

It follows that as $p \to \infty$

$$S = \Big(1+o(1)\Big)\int_{|x_1|\le \frac{\log^2 n}{n^{1/2}}, |x_2|\le \frac{\log^2 n}{n^{3/2}}} \exp\Big(-(1/2+o(1))4\pi^2\sum_{i=1}^n (x_1+ix_2)^2\Big)dx_1dx_2$$
$$\ge \Big(1+o(1)\Big)\int_{|x_1|\le \frac{\log^2 n}{n^{1/2}}, |x_2|\le \frac{\log^2 n}{n^{3/2}}} \exp\Big(-(1/2+o(1))4\pi^2\sum_{i=1}^n 2(x_1^2+i^2x_2^2)\Big)dx_1dx_2$$
$$\ge \Big(1+o(1)\Big)\int_{|x_1|\le \frac{\log^2 n}{n^{1/2}}} \exp\Big(-(1/2+o(1))8\pi^2 n x_1^2\Big)dx_1 \int_{|x_2|\le \frac{\log^2 n}{n^{3/2}}} \exp\Big(-(1/2+o(1))8\pi^2 n^3 x_2^2/6\Big)dx_2.$$

After changing variables $y_1 = \sqrt{8n\pi^2}x_1$ and $y_2 = \sqrt{4\pi^2 n^3/6}x_2$, and using the Gaussian identity $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-y^2/2)dy = 1$, we have $S \ge \Omega(n^{-2})$, completing the proof.

□

To complete the picture, we prove Claim 6.1. In fact, more is true:

Claim A.5. *Let $u_i = (1,i)$ and $v_i = (1,(-1)^{i-1}i)$, $0 \le i \le n$. Assume that $\xi_0,\dots,\xi_n$ are iid copies of a random variable of variance one and bounded $(2+\varepsilon)$-moment. Then we have*

- $\sup_{a\in\mathbb{R}^2}\mathbf{P}(\sum_i \xi_i u_i = a) = O(n^{-2})$;
- $\sup_{a\in\mathbb{R}^2}\mathbf{P}(\sum_i \xi_i v_i = a) = O(n^{-2})$.

Proof of Claim A.5. As the proof for the second estimate is similar, it suffices the prove the first one. We'll use [25]. Assume otherwise that $\rho := \sup_{a\in\mathbb{R}^2}\mathbf{P}(\sum_i \xi_i u_i = a) \ge C_1 n^{-2}$ for some sufficiently large C_1. Then by [25, Theorem 2.5] most of the v_i belongs to a symmetric GAP Q of rank r and size

$$|Q| \le O(\rho^{-1}/n^{r/2}).$$

As Q is symmetric, the difference set $2Q = Q - Q = Q + Q$ contains two linearly independent vectors $u_1 - 0 = (1,1)$ and $u_2 - u_1 = (0,1)$. As such, the rank r of $2Q$ (and also of Q) is at least 2. Plugging this information into the upper bound of $|Q|$, we obtain $|Q| \le O(C_1^{-1}n)$, which is smaller than $n/2$ if C_1 is sufficiently large, a contradiction to the fact that Q contains most of the u_i. □

Appendix B. A remark on Theorem 4.7

B.1. **Deduction of Theorem 4.7 from** [25]**.** For the reader's convenience, we insert here an almost complete treatment, following the proof of [25, Theorem 2.9].

Given a real number w and a random variable z with uniform distribution of parameter N, we define the z-norm of w by

$$\|w\|_z := (\mathbf{E}\|w(z_1 - z_2)\|^2)^{1/2},$$

where z_1, z_2 are two iid copies of z, and $\|x\|$ is the distance of x to the nearest integer.

Fourier analysis. Our first step is to upper bound the small ball probability

Lemma B.2.

$$\rho = \sup_{a\in\mathbb{R}} \mathbf{P}\left(|\sum_i \xi_i v_i - a| \le \beta \right) \le \exp(\pi\beta^2) \int_{\mathbb{R}} \exp(-\sum_{i=1}^n \|\langle v_i, \xi\rangle\|_z^2/2 - \pi\|\xi\|_2^2) d\xi.$$

Next, it is more convenient to consider the scaling $V_\beta := \beta^{-1} \cdot V = \{\beta^{-1}v_1, \dots, \beta^{-1}v_n\}$.

Set $M := 2A \log n$ where A is large enough. From Lemma B.2 and the fact that $\rho \ge n^{-O(1)}$ we easily obtain

$$\int_{|\xi|_2 \le M} \exp(-\frac{1}{2}\sum_{v\in V_\beta} \|v\xi\|_z^2 - \pi|\xi|_2^2) d\xi \ge \frac{\rho}{2}. \tag{43}$$

Large level sets. For each integer $0 \le m \le M$ we define the level set

$$S_m := \left\{ \xi \in \mathbb{R}^d : \sum_{v\in V_\beta} \|\langle v, \xi\rangle\|_z^2 + \|\xi\|_2^2 \le m \right\}.$$

Then it follows from (43) that $\sum_{m\le M} \mu(S_m) \exp(-\frac{m}{2} + 1) \ge \rho$, where $\mu(.)$ denotes the Lebesgue measure of a measurable set. Hence there exists $m \le M$ such that $\mu(S_m) \ge \rho \exp(\frac{m}{4} - 2)$.

Next, since $S_m \subset B(0, \sqrt{m})$, by the pigeon-hole principle there exists a ball $B(x, \frac{1}{2}) \subset B(0, \sqrt{m})$ such that

$$\mu(B(x, \frac{1}{2}) \cap S_m) \ge c_1 \mu(S_m) m^{-1/2} \ge c_1 \rho \exp(\frac{m}{4} - 2) m^{-1/2}.$$

Consider $\xi_1, \xi_2 \in B(x, 1/2) \cap S_m$. By the Cauchy-Schwarz inequality,

$$\sum_{v\in V_\beta} \|\langle v, (\xi_1 - \xi_2)\rangle\|_z^2 \le 4m.$$

Since $\xi_1 - \xi_2 \in B(0, 1)$ and $\mu(B(x, \frac{1}{2}) \cap S_m - B(x, \frac{1}{2}) \cap S_m) \ge \mu(B(x, \frac{1}{2}) \cap S_m)$, if we put

$$T := \{\xi \in B(0, 1), \sum_{i=1}^n \|\langle \xi, v_i\rangle\|_z^2 \le 4m\},$$

then

$$\mu(T) \ge c_1 \rho \exp(\frac{m}{4} - 2) m^{-1/2}.$$

Discretization. Choose N to be a sufficiently large prime (depending on the set T). Define the discrete box

$$B_0 := \{k_1/N : k_1 \in \mathbf{Z}, -N \leq k_1 \leq N\}.$$

We consider all the shifted boxes $x + B_0$, where $x \in [0, 1/N]$. By the pigeon-hole principle, there exists x_0 such that the size of the discrete set $(x_0+B_0)\cap T$ is at least the expectation, $|(x_0+B_0)\cap T| \geq N\mu(T)$.

Let us fix some $\xi_0 \in (x_0 + B_0) \cap T$. Then for any $\xi \in (x_0 + B_0) \cap T$ we have

$$\sum_{v\in V_\beta} \|v(\xi_0 - \xi)\|_z^2 \leq 2\left(\sum_{v\in V_\beta} \|v\xi\|_z^2 + \sum_{v\in V_\beta} \|v\xi_0\|_z^2\right) \leq 16m.$$

Notice that $\xi_0 - \xi \in B_1 := B_0 - B_0 = \{k_1/N, : k_1 \in \mathbf{Z}, -2N \leq k_1 \leq 2N\}$. Thus there exists a subset S of size at least $c_1 N\rho \exp(\frac{m}{4} - 2)m^{-1/2}$ of B_1 such that the following holds for any $s \in S$

$$\sum_{v\in V_\beta} \|vs\|_z^2 \leq 16m.$$

Double counting. We let $y = z_1 - z_2$, where z_1, z_2 are iid copies of z. By definition of S, we have

$$\sum_{s\in S}\sum_{v\in V_\beta} \|vs\|_z^2 \leq 16m|S|$$
$$\mathbf{E}_y \sum_{s\in S}\sum_{v\in V_\beta} \|y(vs)\|_{\mathbb{R}/\mathbf{Z}}^2 \leq 16m|S|.$$

By the uniform property of z, there exist $|y_0| \asymp 1$ such that

$$\sum_{s\in S}\sum_{v\in V_\beta} \|y_0(vs)\|_{\mathbb{R}/\mathbf{Z}}^2 \leq 32m|S|.$$

Let n' be any number between n^ε and n. We say that $v \in V_\beta$ is *bad* if

$$\sum_{s\in S} \|y_0(vs)\|_{\mathbb{R}/\mathbf{Z}}^2 \geq \frac{32m|S|}{n'}.$$

Then the number of bad vectors is at most n'. Let V'_β be the set of remaining vectors. Thus V'_β contains at least $n - n'$ elements.

Dual sets. Consider an arbitrary $v \in V'_\beta$, we have $\sum_{s\in S} \|y_0(sv)\|_{\mathbb{R}/\mathbf{Z}}^2 \leq 32m|S|/n'$.

Set $k := \sqrt{\frac{n'}{64\pi^2 m}}$ and let $V''_\beta := k(V'_\beta \cup \{0\})$. By Cauchy-Schwarz inequality, for any $a \in V''_\beta$ we have

$$\sum_{s\in S} 2\pi^2 \|s(y_0 a)\|^2_{\mathbb{R}/\mathbf{Z}} \le \frac{|S|}{2},$$

which implies

$$\sum_{s\in S} \cos(2\pi s y_0 a) \ge \frac{|S|}{2}.$$

Observe that for any $x \in C(0, \frac{1}{256})$ (the interval of radius 1/256 centered at origin) and any $s \in S \subset C(0,2)$ we always have $\cos(2\pi sx) \ge 1/2$ and $\sin(2\pi sx) \le 1/12$. Thus for any $x \in C(0, \frac{1}{256})$,

$$\sum_{s\in S} \cos\left(2\pi s(y_0 a + x)\right) \ge \frac{|S|}{4} - \frac{|S|}{12} = \frac{|S|}{6}.$$

On the other hand,

$$\begin{aligned}\int_{x\in[0,N]} \left(\sum_{s\in S} \cos(2\pi sx)\right)^2 dx &\le \sum_{s_1,s_2\in S} \int_{x\in[0,N]} \exp\left(2\pi\sqrt{-1}(s_1-s_2)x\right) dx \\ &\ll |S|N.\end{aligned}$$

Hence we deduce the following bound for the Lebesgue measure

$$\mu_{x\in[0,N]}\left((\sum_{s\in S}\cos(2\pi sx))^2 \ge (\frac{|S|}{6})^2\right) \ll_d \frac{|S|N}{(|S|/6)^2} \ll \frac{N}{|S|}.$$

Now use the fact that S has large size, $|S| \gg_N \rho\exp(\frac{m}{4}-2)m^{-1/2}$, and N was chosen to be large enough so that $y_0 V''_\beta + C(0,\frac{1}{256}) \subset [0,N]$, we have

$$\mu(y_0 V''_\beta + C(0,\frac{1}{256})) \ll \rho^{-1}\exp(-\frac{m}{4}+2)m^{1/2}.$$

Thus, we obtain

$$\mu\left(k(V'_\beta\cup\{0\}) + C(0,\frac{1}{256y_0})\right) = O(\rho^{-1}y_0^{-1}\exp(-\frac{m}{4}+2)m^{1/2}). \tag{44}$$

Notice that $k = \sqrt{n'/64\pi^2 m}$, and as n is sufficiently large, $\ell_0 = \sqrt{n'/\log^2 n} \leq k$. Thus we also have

$$\mu\left(\ell_0(V'_\beta \cup \{0\}) + C(0, \frac{1}{256 y_0})\right) = O(\rho^{-1} y_0^{-1} \exp(-\frac{m}{4} + 2) m^{1/2}). \tag{45}$$

Let $D := 1024 y_0$. We approximate each vector v' of V'_β by a closest number in $\frac{\mathbf{Z}}{D\ell_0}$,

$$|v' - \frac{a}{D\ell_0}| \leq \frac{1}{D\ell_0}, \text{ with } a \in \mathbf{Z}.$$

Let A_β be the collection of all such a. It follows from (45) that

$$|\ell_0(A_\beta + C_0(0,1))| = O(\rho^{-1}\ell_0 \exp(-\frac{m}{4} + 2) m^{1/2}) = O(\rho^{-1}\ell_0),$$

where $C_0(0,r)$ is the discrete interval $\{z_1 \in \mathbf{Z}^d : |z_1| \leq r\}$.

Now we apply Theorem [30, Theorem 1.21], it is implied that the set $\ell_0(A_\beta + C_0(0,1))$ belongs to a GAP Q_0 of size $O(\rho^{-1}\ell_0)$, where the implied constant depends on ξ, β and C'. Next, by iterating [29, Theorem 3.40] if necessary, one can assume that Q_0 is C-proper, while the size $|Q_0|$ remains $O(\rho^{-1}\ell_0)$ but the implied constant now also depends on C.

In the next step, one applies [25, Lemma A.2] to "divide" the GAP Q_0, obtaining a C-proper GAP $P \subset \mathbf{Z}$ containing $A_\beta + C_0(0,1)$, which has small rank $1 \leq r = O(1)$, and small size $|P| = O(\rho^{-1}\ell_0\ell_0^{-r}) = O(\rho^{-1})$.

In summary, setting $Q := \frac{\beta}{\ell_0 D} \cdot P$, then the following holds:

- Q has small rank, $r = O(1)$, and small cardinality $|Q| = O(\rho^{-1}\ell_0^{1-r}) = O(\rho^{-1})$;
- for all but at most n' elements a of $\{a_1, \dots, a_n\}$, there exists $q \in Q$ such that

$$|q - a| \leq \beta/D\ell_0;$$

- Q is C-proper;
- as $C_0(0,1) \in P$, there exist $|k_1| \leq L_1, \dots, |k_d| \leq L_d$ such that

$$\beta/D\ell_0 = \sum_i k_i g_i;$$

- as $P \subset \mathbf{Z}$, all steps g_i of Q are *integral multiples* of $\beta/D\ell_0$.

References

[1] A. Bloch and G. Pólya, On the roots of certain algebraic equations, *Proc. London Math. Soc.* **33**(1932), 102-114.

[2] L. Blum, F. Cucker, M. Shub and S. Smale, *Complexity and Real Computation*, Springer-Verlag, New York, 1998.

[3] A. T. Bharucha-Reid and M. Sambandham, *Random polynomials, Probability and Mathematical Statistics*. Academic Press, Inc., Orlando, Fla., 1986.

[4] F. Cucker, T. Krick, G. Malajovich and M. Wschebor, A Numerical Algorithm for Zero Counting. III: Randomization and Condition, *Advances in Applied Mathematics* **48**, 215-248.

[5] A. Edelman and E. Kostlan, How many zeros of a random polynomial are real?, *Bull. Amer. Math. Soc.* (N.S.) **32** (1995), 1–37. Erratum: Bull. Amer. Math. Soc. (N.S.) **33** (1996), 325.
[6] T. Erdélyi, Extensions of the Bloch-Pólya theorem on the number of real zeroes of polynomials, *J. Théor. Nombres Bordeaux* **20** (2008), no. 2, 281-287.
[7] P. Erdős and A. C. Offord, On the number of real roots of a random algebraic equation, *Proc. London Math. Soc.* **6** (1956), 139–160.
[8] K. Farahmand, Topics in random polynomials, *Pitman research notes in mathematics*, series 393. Longman, Harlow, 1998.
[9] I. A. Ibragimov and N. B. Maslova, The average number of zeros of random polynomials, *Vestnik Leningrad. Univ.* **23** (1968), 171–172.
[10] I. A. Ibragimov and N. B. Maslova, The mean number of real zeros of random polynomials. I. Coefficients with zero mean, *Theor. Probability Appl.* **16** (1971), 228–248.
[11] I. A. Ibragimov and N. B. Maslova, The mean number of real zeros of random polynomials. II. Coefficients with a nonzero mean., *Theor. Probability Appl.* **16** (1971), 485–493.
[12] I. A. Ibragimov and N. B. Maslova, The average number of real roots of random polynomials, *Soviet Math. Dokl.* **12** (1971), 1004–1008.
[13] M. Kac, On the average number of real roots of a random algebraic equation, *Bull. Amer. Math. Soc.* **49** (1943) 314–320.
[14] M. Kac, On the average number of real roots of a random algebraic equation. II. *Proc. London Math. Soc.* **50**, (1949), 390–408.
[15] M. Kac, Probability and related topics in physical sciences, *Lectures in Applied Mathematics, Proceedings of the Summer Seminar*, Boulder, Colo., 1957, Vol. I Interscience Publishers, London-New York, 1959.
[16] E. Kostlan, On the distribution of roots of random polynomials, Chapter 38, *From Topology to Computation: Proceeding of the Samefest*, edited by M. W. Hirsch, J.E. , Marsden and M. Shub, Springer-Verlag, NY 1993.
[17] J. E. Littlewood and A. C. Offord, On the distribution of the zeros and a-values of a random integral function. I., *J. Lond. Math. Soc.*, **20** (1945), 120-136.
[18] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation. II. *Proc. Cambridge Philos. Soc.* **35**, (1939), 133-148.
[19] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation. III. *Rec. Math. [Mat. Sbornik] N.S.* **54**, (1943), 277-286.
[20] J. E. Littlewood and A. C. Offord, On the distribution of the zeros and values of a random integral function. II., *Ann. Math.* **49** (1948), 885–952. Errata, **50** (1949), 990-991. 976), 35–58.
[21] B. F. Logan and L. A. Shepp, Real zeros of random polynomials. *Proc. London Math. Soc.* **18** (1968), 29–35.
[22] B. F. Logan and L. A. Shepp, *Real zeros of random polynomials. II. Proc. London Math. Soc.* **18** (1968), 308–314.
[23] N. B. Maslova, The variance of the number of real roots of random polynomials. *Teor. Vero- jatnost. i Primenen.* **19** (1974), 36–51.
[24] N. B. Maslova, The distribution of the number of real roots of random polynomials. *Theor. Probability Appl.* **19** (1974), 461–473
[25] H. Nguyen and V. Vu, Optimal Inverse Littlewood-Offord theorems, *Adv. Math.* **226** (2011), no. 6, 5298–5319.
[26] H. Nguyen, O. Nguyen and V. Vu, On the number of real roots of random polynomials, *submitted*.
[27] R. Peled, A. Sen and O. Zeitouni, Double roots of random Littlewood polynomials, `http://arxiv.org/abs/1409.2034`.
[28] D.C. Stevens, The average number of real zeros of a random polynomial. *Comm. Pure Appl. Math.* **22** (1969), 457–477.
[29] T. Tao and V. Vu, Additive Combinatorics, Cambridge Univ. Press, 2006.
[30] T. Tao and V. Vu, John-type theorems for generalized arithmetic progressions and iterated sumsets, *Adv. Math.* 219 (2008), no. 2, 428-449.
[31] T. Tao and V. Vu, Local universality of zeros of random polynomials, *to appear in IMRN*.
[32] I. Todhunter, A history of the mathematical theory of probability, Stechert, New York, 1931.
[33] Y. Wang, Bounds on the average number of real roots of a random algebraic equation, *Chinese Ann. Math.* Ser. A 4 (1983), 601-605.
[34] J. E. Wilkins, An asymptotic expansion for the expected number of real zeros of a random polynomial, Proc. Amer. Math. Soc. 103 (1988), 1249-1258.

E-mail address: `yen.do@virginia.edu`

Department of Mathematics, The University of Virginia, Charlottesville, VA 22904

E-mail address: nguyen.1261@math.osu.edu

Department of Mathematics, The Ohio State University, Columbus, OH 43210

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

Department of Mathematics, Yale University, New Haven, CT 06511