EXPONENTIAL CONCENTRATION FOR THE NUMBER OF ROOTS OF RANDOM TRIGONOMETRIC POLYNOMIALS

HOI H. NGUYEN AND OFER ZEITOUNI

Abstract. We show that the number of real roots of random trigonometric polynomials with i.i.d. coefficients, which are either bounded or satisfy the logarithmic Sobolev inequality, satisfies an exponential concentration of measure.

1. Introduction

Consider a random trigonometric polynomial of degree n

$$P_n(x) = \frac{1}{\sqrt{n}} \sum_{k=1}^n a_k \cos(kx) + b_k \sin(kx), \tag{1}$$

where a_k, b_k are i.i.d. copies of a random variable ξ of mean zero and variance one. Let N_n denote the number of roots of $P_n(x)$ for $x \in [-\pi, \pi]$. It is known from a work of Qualls [27] that when ξ is standard gaussian then

$$\mathbf{E} N_n = 2\sqrt{(2n+1)(n+1)/6}.$$

By a delicate method based on the Kac-Rice formula, about ten years ago Granville and Wigman showed the following.

Theorem 1.1 ([15]). *When ξ is standard gaussian, there exists an explicit constant $c_{\mathbf{g}}$ such that*

$$\mathrm{Var}(N_n) = (c_{\mathbf{g}} + o(1))n.$$

Furthermore,

$$\frac{N_n - \mathbf{E}N_n}{\sqrt{c_{\mathbf{g}} n}} \xrightarrow{d} \mathbf{N}(0,1).$$

This confirms a heuristic by Bogomolny, Bohigas and Leboeuf. More recently, Azaïs and León [4] provided an alternative approach based on the Wiener chaos decomposition. They showed that $Y_n(t) = P_n(t/n)$ converges in certain strong sense to the stationary gaussian process $Y(t)$ of covariance $r(t) = \sin(t)/t$, from which variance and CLT can be deduced.

These methods do not seem to work for other ensembles of ξ. Under a more general assumption, a recent result by O. Nguyen and Vu shows the following. Throughout, we use the standard $O(\cdot)$ notation, see Section 1.6 for definition.

The first author is supported by National Science Foundation CAREER grant DMS-1752345. The second author is partially supported by a US-Israel BSF grant. This work was initiated when both authors visited the American Institute of Mathematics in August 2019. We thank AIM for its hospitality.

Theorem 1.2 ([24]). *Assume that ξ has a bounded $(2+\varepsilon_0)$-moment for a positive constant ε_0. Then, there exists a constant $c>0$ such that*

$$\mathbf{E}N_n = (2/\sqrt{3} + O(n^{-c}))n$$

and [1]

$$\mathrm{Var}(N_n) = O(n^{2-c}).$$

Furthermore, assuming that $|\xi|$ has finite moments of all order, under an anti-concentration estimate on ξ of the form that there exists an $r>0$ and $a \in \mathbb{R}$ for which $\mathbf{P}(\xi \in A) \geq c\mathrm{Leb}(A)$ for all $A \subset B(a,r)$, a special case of a recent result by Bally, Caramellino, and Poly [7] regarding the number $N_n([0,\pi])$ of roots over $[0,\pi]$ [2] reads as follows.

Theorem 1.3 ([7]). *There exists a constant $c'_{\mathbf{g}}$ such that*

$$\lim_{n\to\infty} \frac{1}{n}\mathrm{Var}(N_n([0,\pi])) = c'_{\mathbf{g}} + \frac{1}{30}\mathbf{E}(\xi^4 - 3).$$

This result has been generalized to more general ξ in [10]. Our goal in this note is rather different from the results above, in that we are interested in the *concentration* (deviation) of N_n rather than the asymptotic statistics. In some way, our work is motivated by a result by Nazarov and Sodin [21] on the concentration of the number of nodal domains of random spherical harmonics, and by the exponential concentration phenomenon of the number of zeros of stationary gaussian process [5]. See also [14]. We will show the following.

Theorem 1.4. *Let C_0 be a given positive constant, and suppose that either $|\xi|$ is bounded almost surely by C_0, or that its law satisfies the logarithmic Sobolev inequality* (5) *with parameter C_0. Then there exist constants c, c' such that for $\varepsilon \geq n^{-c}$ we have that*

$$\mathbf{P}(|N_n - \mathbf{E}N_n| \geq \varepsilon n) = O(e^{-c'\varepsilon^9 n}).$$

The above result immediately implies that $\frac{N_n}{n} \to \frac{2}{\sqrt{3}}$ almost surely. Note that the latter result has also been obtained recently by Angst and Poly ([2, Theorem 6]) where the authors there used the moment method to establish a polynomial-type speed of convergence [3].

In case ξ is Gaussian, Theorem 1.4 bears resemblance to [5]. Note however that it is not immediate to read Theorem 1.4 from [5], since there is no direct relation between the length of the time interval T in the latter and n. It is plausible that with some effort, one could modify the proof technique in [5] to cover this case. Our methods however are completely different and apply in particular to the Bernoulli case.

We also remark that in the Gaussian case, by following [11] our result yields the following equidistribution interpretation. Consider the curve $\gamma(x)$ on the unit sphere S^{2n-1} defined by our polynomial,

$$\gamma(x) = \frac{1}{\sqrt{n}}\big(\cos(x), \sin(x), \dots, \cos(nx), \sin(nx)\big), x \in [-\pi, \pi].$$

For each x, let $\gamma(x)_\perp$ be the set (known as "great hypercircles") of vectors on S^{2n-1} that are orthogonal to $\gamma(x)$. Let $\gamma_\perp$ be the region (counting multiplicities) swept by $\gamma(x)_\perp$ when x varies in $[-\pi,\pi]$. Then $\gamma_\perp$ covers S^{2n-1} *uniformly* in the sense that the Haar measure of those sphere

[1] See [10, Section 8].

[2] We remark that the authors of [7] work with roots over $[0,\pi]$.

[3] The assumption of ξ in [2], on the other hand, is slightly more general.

points that are covered k-times, where $k \notin [(2/\sqrt{3}-\varepsilon)n, (2/\sqrt{3}+\varepsilon)n]$, is at most $e^{-c'\varepsilon^9 n}$ whenever $n^{-c} \le \varepsilon$. In another direction, our result also implies an exponential-type upper bound for the persistence probability that $P_n(x)$ does not have any root (over $[-\pi,\pi]$, and hence entirely).

Our overall method is somewhat similar to [21], but the situation for trigonometric functions seems to be rather different compared to spherical harmonics, for instance we don't seem to have the analogs of [21, Claim 2.2] or [21, Claim 2.4] for trigonometric polynomials. Another different aspect of our work is its universality, that the concentration phenomenon holds for many other ensembles where we clearly don't have any invariance property at hands. One of the main ingredients in the proof is the phenomenon of root repulsion, which has also been recently studied in various ensembles of random polynomials, see [9, 13, 23, 26] among others. Our method is robust, and seems to be applicable to other models of random polynomials. In any case, it remains an interesting problem to optimize the range of c in Theorem 1.4. Although our proofs give explicit values of c, c', they are far from being optimal.

Remark 1.5. *Versions of Theorem 1.4 hold for other types of ξ, which are neither bounded nor satisfy the logarithmic Sobolev inequality. Specifically, the following holds for $\varepsilon \ge n^{-c}$, for some sufficiently small constant $c > 0$.*

(i) Assume $|\xi|$ has sub-exponential tail. Then, for some positive constant δ depending on the tail of $|\xi|$, we have the sub-exponential concentration

$$\mathbf{P}(|N_n - \mathbf{E}N_n| \ge \varepsilon n) = O(e^{-(\varepsilon n)^\delta}).$$

(ii) Fix $C > 0$. Assume that is $\mathbf{E}(|\xi|^{C'}) < \infty$ for some sufficiently large $C' = C'(C)$, then we have the polynomial concentration

$$\mathbf{P}(|N_n - \mathbf{E}N_n| \ge \varepsilon n) = O((\varepsilon n)^{-C}).$$

Both cases can be established by taking $C_0 = n^{c_0}$ in Theorem 2.5 below, with an appropriate c_0. We refer the reader to the end of Section 6 for further details.

Before concluding this section we record here a corollary of Theorem 1.2 which will be useful later: for ξ as in the theorem, for any $\varepsilon > 0$ we have

$$\mathbf{P}(|N_n - \mathbf{E}N_n| \ge \varepsilon n/2) = O(\varepsilon^{-2} n^{-c}). \tag{2}$$

1.6. **Notation.** We will assume $n \to \infty$ throughout the paper. We write $X = O(Y)$, $X \ll Y$, or $Y \gg X$ if $|X| \le CY$ for some absolute constant C. If C depends on another parameter τ, we will write $Y = O_\tau(X)$. We write $X \asymp Y$ if $X \gg Y$ and $Y \gg X$. Also, if $Y_n/|X_n| \to \infty$ (as $n \to \infty$) then we write $X_n = o(Y_n)$. In what follows, if not specified otherwise, all of the norms on Euclidean spaces are L_2-norm (i.e. $d_2(.)$ distance). Finally, we use Leb to denote the Lebesgue measure on $\mathbf{T}$.

2. Some supporting lemmas and the proof method

In this section we gather several well-known results regarding trigonometric polynomials. We then describe in Section 2.7 our proof strategy for Theorem 1.4.

On the deterministic side, a useful ingredient is the classical Bernstein's inequality in $L_2(\mathbf{T})$, where $\mathbf{T} = [-\pi, \pi]$. The proof is immediate from the orthogonality relations satisfied by the trigonometric base.

Theorem 2.1. *Let $f(x) = \sum_{k=0}^{n} a_k \cos(kx) + b_k \sin(kx), x \in \mathbf{T}$. Then,*

$$\int_{x\in\mathbf{T}} (f'(x))^2 dx \le n^2 \int_{x\in\mathbf{T}} f(x)^2 dx.$$

Another crucial inequality we will be using is the so-called large sieve inequality.

Theorem 2.2. [16, Theorem 7.7][20, (1.1)] *Assume that f is as in Theorem 2.1. Then for any $-\pi \le x_1 < x_2 < \cdots < x_M \le \pi$ we have*

$$\sum_{i=1}^{M} |f(x_i)|^2 \le \frac{2n+\delta^{-1}}{2\pi} \int_{x\in\mathbf{T}} f(x)^2 dx,$$

where δ is the minimum of the gaps between x_i, x_{i+1} on the torus.

As a corollary, we obtain the following.

Corollary 2.3. *Assume that $\|f\|_{L_2(\mathbf{T})} \le \tau$. Then the set of $x \in \mathbf{T}$ with $|f(x)| \ge \lambda$ or $|f'(x)| \ge \lambda n$ is contained in the union of $2M$ intervals of length 2δ, where $M \le \frac{2n+\delta^{-1}}{2\pi}\frac{\tau^2}{\lambda^2}$.*

Proof. Choose a maximal set of δ-separated points x_i for which $|f(x_i)| \ge \lambda$. Then by Theorem 2.2 we have $M\lambda^2 \le \frac{2n+\delta^{-1}}{2\pi}\tau^2$. We can apply the same argument for f' where by Bernstein's inequality we have $\|f'\|_2 \le n\|f\|_2 \le n\tau$. □

We next introduce an elementary interpolation result (see for instance [8, Section 1.1, E.7]).

Lemma 2.4. *Assume that a trigonometric polynomial P_n has at least m zeros (counting multiplicities) in an interval I of length r. Then*

$$\max_{\theta\in I} |P_n(\theta)| \le (\frac{4er}{m})^m \max_{x\in I} |P_n^{(m)}(x)|$$

as well as

$$\max_{\theta\in I} |P_n'(\theta)| \le (\frac{4er}{m-1})^{m-1} \max_{x\in I} |P_n^{(m)}(x)|.$$

Consequently, if P_n has at least m roots on an interval I with length smaller than $(1/8e)m/n$, then for any interval I' of length $(1/8e)m/n$ and $I \subset I'$ we have

$$\max_{\theta\in I'} |P_n(\theta)| \le (\frac{1}{2})^m (\frac{1}{n})^m \max_{x\in I'} |P_n^{(m)}(x)| \tag{3}$$

as well as

$$\max_{\theta\in I'} |P_n'(\theta)| \le n \times (\frac{1}{2})^{m-1} (\frac{1}{n})^m \max_{x\in I'} |P_n^{(m)}(x)|. \tag{4}$$

Proof. It suffices to show the estimates for P_n because P_n' has at least $m-1$ roots in I. For P_n, by Hermite interpolation using the roots x_i we have that for any $\theta \in I$ there exists $x \in I$ so that

$$|P_n(\theta)| = |\frac{P_n^{(m)}(x)}{m!} \prod_i (\theta - x_i)| \le \max_{x\in I} |P_n^{(m)}(x)| \frac{r^m}{m!}.$$

□

On the probability side, for bounded random variables we will rely on the following consequence of McDiarmid's inequality.

Theorem 2.5. *Assume that* $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$*, where* ξ_i *are iid copies of* ξ *of mean zero, variance one, taking values in* $\Omega = [-C_0, C_0]$*. Let* $\mathcal{A}$ *be a set of* Ω^n*. Then for any* $t > 0$ *we have*

$$\mathbf{P}(\boldsymbol{\xi} \in \mathcal{A})\mathbf{P}(d_2(\boldsymbol{\xi}, \mathcal{A}) \geq t\sqrt{n}) \leq 4\exp(-t^4 n/16C_0^4).$$

For random variables ξ satisfying the log-Sobolev inequality, that is so that there is a positive constant C_0 such that for any smooth, bounded, compactly supported functions f we have

$$\mathrm{Ent}_\xi(f^2) \leq C_0 \mathbf{E}_\xi |\nabla f|^2, \tag{5}$$

where $\mathrm{Ent}_\xi(f) = \mathbf{E}_\xi(f\log f) - (\mathbf{E}_\xi(f))(\log \mathbf{E}_\xi(f))$, we use the following.

Theorem 2.6. *Assume that* $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$*, where* ξ_i *are iid copies of* ξ *satisfying* (5) *with a given* C_0*. Let* $\mathcal{A}$ *be a set in* $\mathbb{R}^n$*. Then for any* $t > 0$ *we have*

$$\mathbf{P}\big(d_2(\boldsymbol{\xi}, \mathcal{A}) \geq t\sqrt{n}\big) \leq 2\exp\big(-\mathbf{P}^2(\boldsymbol{\xi} \in \mathcal{A})t^2 n/4C_0\big).$$

In particularly, if $\mathbf{P}(\boldsymbol{\xi} \in \mathcal{A}) \geq 1/2$ *then* $\mathbf{P}(d_2(\boldsymbol{\xi}, \mathcal{A}) \geq t\sqrt{n}) \leq 2\exp(-t^2 n/16C_0)$*. Similarly if* $\mathbf{P}(d_2(\boldsymbol{\xi}, \mathcal{A}) \geq t\sqrt{n}) \geq 1/2$ *then* $\mathbf{P}(\boldsymbol{\xi} \in \mathcal{A}) \leq 2\exp(-t^2 n/16C_0)$*.*

For consistency with Theorem 2.5, we sometimes use Ω^n in place of $\mathbb{R}^n$ in various applications of Theorem 2.6.

For completeness, the proofs of these well-known results will be presented in Section A.

2.7. **Proof strategy.** We now sketch our proof method, which broadly speaking, follows the perturbation framework of [21]. Our starting point is the input (2) that $N_n(P_n)$ is moderately concentrated around its mean $\mathbf{E}N_n$. This will ensure that various sets have probability at least 1/2 so that the concentration estimates in Theorems 2.5 and 2.6 will give effective bounds, see for an example the proof of Corollary 5.3.

- (i) Our first highlight is Section 4 where we show that it is highly unlikely that there is a large set of *unstable* intervals where $|P_n|$ and $|P_n'|$ are both small (see Definition 4.1 for this notion). This is justified by relying on a repulsion estimate (Theorem 3.1), on Theorem 2.2, and also on Theorem 2.5 and 2.6 to exploit exponential concentration. We call polynomials in this unlikely set *exceptional*.
- (ii) Additionally, we will show in Corollary 5.3 of Section 5 that the number of roots over the unstable intervals is small for any non-exceptional P_n. We justify this deterministic result by using the elementary tool of Lemma 2.4.
- (iii) Building on these results, and with the stability result of Corollary 6.2, if P_n is not too exceptional and $N_n(P_n)$ is close to $\mathbf{E}N_n$ then $N_n(Q_n)$ is also close to $\mathbf{E}N_n$, as long as $\|P_n - Q_n\|_2$ is small. As such, geometric concentration tools such as Theorem 2.5 and 2.6 can be invoked once more to show that indeed N_n satisfies near exponential concentration, as desired.

3. Repulsion estimate

We show that the measure of $t \in [-\pi, \pi]$ where both $|P_n(t)|$ and $|P_n'(t)|$ are small is negligible. More precisely we will be working with the following condition.

Condition 1. *Let $0 < \tau \leq 1/64$ be given, and let C_0' be a positive constant to be chosen sufficiently large. Assume that $t \in [-\pi, \pi]$ is such that there do not exist integers k with $|k| \leq C_0'$ satisfying*

$$\|kt/\pi\|_{\mathbb{R}/\mathbb{Z}} \leq n^{-1+8\tau}.$$

Here $\|.\|_{\mathbb{R}/\mathbb{Z}}$ is the distance to the nearest integer.

Theorem 3.1. [10, Theorem 2.1] *Assume that ξ has mean zero and variance one. Then as long as $\alpha > 1/n$, $\beta > 1/n$ and t satisfies Condition 1 with given τ, C_0' we have*

$$\mathbf{P}\big(|P_n(t)| \leq \alpha \wedge |P_n'(t)| \leq \beta n\big) = O_{\tau, C_0'}(\alpha\beta).$$

To see that Theorem 3.1 can indeed be deduced from [10], note that we can view the event in Theorem 3.1 as a random walk event in $\mathbb{R}^2$

$$\frac{1}{\sqrt{2n}} \sum_{i=1}^{n} (\xi_i \mathbf{v}_i + \xi_i' \mathbf{v}_i') \in [-\alpha, \alpha] \times [-\beta, \beta],$$

where ξ_i, ξ_i' are iid copies of the random variables ξ, with

$$\mathbf{v}_i := (\cos(it), -\frac{i}{n}\sin(it)) \text{ and } \mathbf{v}_i' := (\sin(it), \frac{i}{n}\cos(it)).$$

Remark 3.2. *The full version of* [10, Theorem 2.1] *allows for $\alpha, \beta > n^{-C}$ (for any given $C > 0$). As we only need here $C = 1$, a simpler and more direct proof can be given. We refer the reader to the preprint version of this article [25], where the argument is detailed.*

4. Exceptional polynomials are rare

This current section is motivated by the treatment in [21, Section 4.2]. We begin with a definition. Let $R > 0$ be a sufficiently large constant. Cover $\mathbf{T}$ by $\frac{2\pi n}{R}$ open interval I_i of length (approximately) R/n each. Let $3I_i$ be the interval of length $3R/n$ having the same midpoint with I_i.

Definition 4.1. Fix $\alpha, \beta, \delta > 0$ and R as above. We call an interval I_i *stable* for a function f if there is no point in $x \in 3I_i$ such that both $|f(x)| \leq \alpha$ and $|f'(x)| \leq \beta n$. We call f *exceptional* if the number of unstable intervals is at least δn. We call f not exceptional otherwise.

Note that the notion of exceptional f depends on the parameters $\alpha, \beta, \delta, R, n$, but we will not emphasize it in the notation.

For convenience, for each $P_n(x) = \frac{1}{\sqrt{n}} \sum_{k=1}^n a_k \cos(kx) + b_k \sin(kx)$ we assign a unique (unscaled) vector $\mathbf{v}_{P_n} = (a_1, \dots, a_n, b_1, \dots, b_n)$ in Ω^{2n}, which is a random vector when P_n is random.

Let $\mathcal{E}_e = \mathcal{E}_e(R, \alpha, \beta; \delta)$ denote the set of vectors $\mathbf{v}_{P_n}$ in Ω^{2n} associated to exceptional polynomials P_n. Our goal in this section is the following.

Theorem 4.2 (Exceptional polynomials are rare)**.** *Assume that α, β, δ satisfy*

$$\alpha \asymp \delta^{3/2}, \beta \asymp \delta^{3/4}, \delta > n^{-2/5}. \tag{6}$$

Assume that ξ is as in Theorem 1.4. Then we have

$$\mathbf{P}\Big(\mathbf{v}_{P_n} \in \mathcal{E}_e\Big) \leq e^{-c\delta^8 n},$$

where c is allowed to depend on C_0 (one can take $c = 1/16C_0^4$ in the setup of Theorem 2.5 and $c = 1/16C_0$ in the setup of Theorem 2.6.)

We now discuss the proof. The first part is deterministic. Assume that f (playing the role of P_n) is exceptional, then there are $K = \lfloor \delta n/3 \rfloor$ unstable intervals that are R/n-separated (and hence $4/n$-separated, as long as R is chosen larger than 4).

Now for each unstable interval in this separated family we choose $x_j \in 3I_j$ where $|f(x_j)| \le \alpha$ and $|f'(x_j)| \le \beta n$ and consider the interval $B(x_j, \gamma/n)$ for some $\gamma < 1$ chosen sufficiently small (given δ). Let

$$M_j := \max_{x \in B(x_j,\gamma/n)} |f''(x)|.$$

By Theorem 2.2 and Theorem 2.1 we have

$$\sum_{j=1}^{K} M_j^2 \le \frac{2n + (4/n)^{-1}}{2\pi} \int_{x\in \mathbf{T}} f''(x)^2 dx \le n^5 \int_{x \in \mathbf{T}} f(x)^2 dx.$$

Now we use randomness: in both the bounded and the log-Sobolev cases we have $\|f\|_2 \ge 2$ with exponentially small probability, so without loss of generality it suffices to assume in what follows that $\|f\|_2 \le 2$. We then infer from the above that the number of j for which $M_j \ge C_2 \delta^{-1/2} n^2$ is at most $2C_2^{-2}\delta n$. Hence for at least $(1/3 - 2C_2^{-2})\delta n$ indices j we must have $M_j < C_2 \delta^{-1/2} n^2$.

Fix x_j as the center of such interval $B(x_j, \gamma/n)$. By Taylor expansion of order two around x_j, we obtain for any x in this interval,

$$|f(x)| \le \alpha + \beta\gamma + C_2 \delta^{-1/2}\gamma^2/2 \quad \text{and} \quad |f'(x)| \le (\beta + C_2\delta^{-1/2}\gamma)n.$$

So far we have dealt with one such exceptional polynomial. We now wish to perturb it and show that an appropriate perturbation is also exceptional. Toward this end, consider a (perturbing) trigonometric polynomial g such that $\|g\|_2 \le \tau$, of the form $g(x) = \frac{1}{\sqrt{n}}(\sum_{k=1}^n a_k' \cos(kx) + b_k' \sin(kx))$, and write $h = f + g$. Then, similarly to Corollary 2.3), as the intervals $B(x_j, \gamma/n)$ are $4/n$-separated, by Theorem 2.2 we have

$$\sum_j \max_{x \in B(x_j,\gamma/n)} g(x)^2 \le 8n\|g\|_2^2 \le 8n\tau^2$$

and

$$\sum_j \max_{x \in B(x_j,\gamma/n)} g'(x)^2 \le 8n\|g'\|_2^2 \le 8n^3\tau^2.$$

Hence, again by an averaging argument, the number of intervals where either $\max_{x\in B(x_j,\gamma/n)} |g(x)| \ge C_3\delta^{-1/2}\tau$ or $\max_{x \in B(x_j,\gamma/n)} |g'(x)| \ge C_3 \delta^{-1/2}\tau n$ is bounded from above by $(1/3 - 2C_2^{-2})\delta n/2$ if C_3 is sufficiently large. On the remaining at least $(1/3 - 2C_2^{-2})\delta n/2$ intervals, we have simultaneously that

$$|h(x)| \le \alpha' \quad \text{and} \quad |h'(x)| \le \beta' n, \tag{7}$$

where

$$\alpha' = \alpha + \beta\gamma + C_2\delta^{-1}\gamma^2/2 + C_3\delta^{-1/2}\tau \text{ and } \beta' = \beta + C_2\delta^{-1/2}\gamma + C_3\delta^{-1/2}\tau.$$

Thus, with Leb the Lebesgue measure and

$$\mathcal{U} = \mathcal{U}(\alpha,\beta,\gamma,\delta,\tau,C_1,C_2,C_3) = \{\mathbf{v}_h \subset \mathbb{R}^{2n} : \text{Leb}(x : (7) \text{ holds}) \ge (1/3 - 2C_2^{-2})\delta\gamma\},$$

the perturbation h has indeed $\mathbf{v}_h \in \mathcal{U}$ (because the set of x in the definition of $\mathcal{U}$ contains $(1/3 - 2C_2^{-2})\delta n/2$ intervals of length $2\gamma/n$). Putting together the above, we have obtained the following claim.

Claim 4.3. *Assume that $\mathbf{v}_{P_n} \in \mathcal{E}_e$. Then for any g with $\|g\|_2 \le \tau$ we have $\mathbf{v}_{P_n+g} \in \mathcal{U}$. In other words,*

$$\left\{\mathbf{v} \in \Omega^{2n}, d_2(\mathcal{E}_e, \mathbf{v}) \le \tau\sqrt{2n}\right\} \subset \mathcal{U}.$$

We next turn to a probabilistic argument. To apply concentration of measure, we will show that $\mathbf{P}(\mathbf{v}_{P_n} \in \mathcal{U}) < 1/2$. Indeed, let $\mathbf{T}_e$ denote the collection of $x \in \mathbf{T}$ which can be $n^{-1+8\tau}$ approximated by rational numbers of bounded height (see Condition 1, here we choose $\tau = 1/64$). Thus $\mathbf{T}_e$ is a union of a bounded number of intervals of length $n^{-1+8\tau}$. For each P_n, let $B(P_n)$ (and $B_e(P_n)$) be the measurable set of $x \in \mathbf{T}$ (or $x \in \mathbf{T}_e^c$ respectively) such that $\{|P_n(x)| \le \alpha'\} \wedge \{|P_n'(x)| \le \beta' n\}$. Then the Lebesgue measure of $B(P_n)$, $\mathrm{Leb}(B(P_n))$, is bounded by $\mathrm{Leb}(B_e(P_n)) + O(n^{-1+8\tau})$, which in turn can be bounded by

$$\mathbf{E}\mathrm{Leb}(B_e(P_n)) = \int_{x\in\mathbf{T}_e^c} \mathbf{P}(\{|P_n(x)| \le \alpha'\} \wedge \{|P_n'(x)| \le n\beta'\})dx = O(\alpha'\beta'),$$

where we used Theorem 3.1 for each x. It thus follows that $\mathbf{E}\mathrm{Leb}(B(P_n)) = O(\alpha'\beta') + O(n^{-1+8\tau})$. So by Markov inequality,

$$\mathbf{P}(\mathbf{v}_{P_n} \in \mathcal{U}) \le \mathbf{P}\big(\mathrm{Leb}(B(P_n)) \ge (1/3 - 2C_2^{-2})\delta\gamma\big) = O(\alpha'\beta'/\delta\gamma) < 1/2 \tag{8}$$

if α, β are as in (6) and then γ, τ are chosen appropriately, for instance as

$$\gamma \asymp \delta^{5/4}, \tau \asymp \delta^2. \tag{9}$$

Proof. (of Theorem 4.2) Let $\mathcal{A} = \Omega^{2n}\backslash\mathcal{U}$. By Claim 4.3, if $\mathbf{v} \in \mathcal{E}_e$ then $d_2(\mathbf{v}, \mathcal{A}) \ge \tau\sqrt{2n}$. Also, by (8), we have $\mathbf{P}(\mathbf{v}_n \in \mathcal{A}) \ge 1/2$. Hence Theorems 2.5 and 2.6 applied to $\mathcal{A}$ imply that

$$\mathbf{P}(\mathbf{v}_n \in \mathcal{E}_e) \le 8e^{-c\tau^4 n},$$

where $c = 1/16C_0^4$ in the case of Theorem 2.5 and $c = 1/16C_0$ in the case of Theorem 2.6. □

5. Roots over unstable intervals

In this section we show the following deterministic lemma.

Lemma 5.1. *Let $\varepsilon < e^{-1}$ be given as in Theorem 1.4. Assume that the parameters α, β, τ are chosen as in (6) and (9). Assume that a trigonometric polynomial P_n has at least $\varepsilon n/2$ roots over δn disjoint intervals of length R/n, where*

$$\delta \le \frac{c_0\varepsilon}{\log(1/\varepsilon)} \tag{10}$$

with $c_0 = 1/1024eR$. Then there is a set $A \subset \mathbf{T}$ of measure at least $\frac{\varepsilon}{1024e}$ on which

$$\max_{x\in A}|f(x)| \le \alpha \text{ and } \max_{x\in A}|f'(x)| \le \beta n.$$

Before proving this result, we deduce a key consequence that non-exceptional polynomials cannot have too many roots over the unstable intervals.

Corollary 5.2. *Let the parameters $\varepsilon, \alpha, \beta, \tau, \delta$ and R be as in Lemma 5.1. Then a non-exceptional P_n cannot have more than $\varepsilon n/2$ roots over any δn intervals I_i from Section 4. In particularly, P_n cannot have more than $\varepsilon n/2$ roots over the unstable intervals.*

Proof. If P_n has more than $\varepsilon n/2$ roots over some δn intervals I_i, then Lemma 5.1 implies the existence of a set $A = A(P_n)$ that intersects with the set of stable intervals so that $\max_{x\in A}|P_n(x)| \le \alpha$ and $\max_{x\in A}|P_n'(x)| \le \beta n$ (because, recall Definition 4.1, the unstable intervals have measure at most δR, and $\varepsilon/(1024e) > \delta R$). However, this is impossible because for any x in the union of the stable intervals we have either $|P_n(x)| > \alpha$ or $|P_n'(x)| > \beta n$. □

We now give an elemantary proof of Lemma 5.1. The main idea is that if P_n has many roots over a small union of intervals, then we can use Hermite interpolation (Lemma 2.4) to show that $|P_n|$ and $|P_n'|$ are small over a set A of non-negligible measure.

Proof of Lemma 5.1. Among the given δn intervals we first throw away those of less than $\varepsilon\delta^{-1}/4$ roots, hence there are at least $\varepsilon n/4$ roots left. For convenience we denote the remaining intervals by $J_1, \dots, J_M$, where $M \le \delta n$, and let $m_1, \dots, m_M$ denote the number of roots over each of these intervals respectively.

In the next step (which is geared towards the use of (3) and (4) of Lemma 2.4), we expand the intervals J_j to larger intervals $\bar{J}_j$ (considered as union of consecutive closed intervals appearing at the beginning of Section 4) of length $\lceil cm_j/R\rceil \times (R/n)$ with $c = 1/(16e)$. Furthermore, if the expanded intervals $\bar{J}_{i_1}', \dots, \bar{J}_{i_k}'$ of $\bar{J}_{i_1}, \dots, \bar{J}_{i_k}$ form an intersecting chain, then we create a longer interval $\bar{J}'$ of length $\lceil c(m_{i_1} + \dots + m_{i_k})/R\rceil \times (R/n)$, which contains them and therefore contains at least $m_{i_1} + \dots + m_{i_k}$ roots. After the merging process, we obtain a collection $\bar{J}_1', \dots, \bar{J}_{M'}'$ with the number of roots $m_1', \dots, m_{M'}'$ respectively, so that $\sum m_i' \ge \varepsilon n/2$. Note that now $\bar{J}_i'$ has length $\lceil cm_i'/R\rceil \times (R/n) \approx cm_i'/n$ (because by assumption $\varepsilon\delta^{-1}$ is sufficiently large compared to R) and the intervals are R/n-separated.

Next, consider the sequence

$$d_l := 2^l \varepsilon\delta^{-1}/4, l \ge 0.$$

We classify the sequence $\{m_i'\}$ into groups G_l of intervals of m_i' roots where

$$d_l \le m_i' < d_{l+1}.$$

Assume that each group G_l has $k_l = |G_l|$ distinct extended intervals. As each of these intervals has between d_l and d_{l+1} roots, we have

$$\sum_l k_l d_l \ge \sum_i m_i'/2 \ge \varepsilon n/8.$$

For given α, β, we call an index l *bad* if

$$(1/2)^{d_l}(n/2k_l)^{1/2} \ge \lambda = \min\{\alpha/4, \beta/4\}.$$

That is when

$$k_l \le \frac{n}{2\lambda^2 4^{d_l}}.$$

The total number of roots over the intervals corresponding to bad indices can be bounded by

$$\sum_{l \text{ bad index}} k_l d_{l+1} \le \frac{n}{2\lambda^2}\sum_{l=0}^{\infty}\frac{2d_l}{4^{d_l}} \le \frac{n}{\lambda^2 2^{\varepsilon\delta^{-1}}} \asymp \frac{n}{\delta^3 2^{\varepsilon\delta^{-1}}} \le \varepsilon n/32$$

where we used the fact that $\delta \le \frac{c_0\varepsilon}{\log(1/\varepsilon)}$ for some small constant c_0.

Now consider a group G_l of *good* index l. Notice that by definition these intervals have length approximately between cd_l/n and $2cd_l/n$. Let I be an interval among the k_l intervals in G_l. By

Lemma 2.4 and by definition we have

$$\max_{x\in I}|P_n(x)| \le (\frac{1}{2})^{d_l}(\frac{1}{n})^{d_l}\max_{x\in I}|P_n^{(d_l)}(x)| \le \frac{\lambda}{(n/2k_l)^{1/2}}(\frac{1}{n})^{d_l}\max_{x\in I}|P_n^{(d_l)}(x)| \tag{11}$$

as well as

$$\max_{x\in I}|P_n'(x)| \le n\times(\frac{1}{2})^{d_l-1}(\frac{1}{n})^{d_l}\max_{x\in I}|P_n^{(d_l)}(x)| \le n\times\frac{2\lambda}{(n/2k_l)^{1/2}}(\frac{1}{n})^{d_l}\max_{x\in I}|P_n^{(d_l)}(x)|, \tag{12}$$

where we used the fact that $(1/2)^{d_l}(n/2k_l)^{1/2} < \lambda$.

On the other hand, as these k_l intervals are R/n-separated (and hence $4/n$-separated), by the large sieve inequality (Theorem 2.2) and by an iterated use of the Bernstein inequality (Theorem 2.1) we have

$$\sum_{\bar{J}_i'\in G_l}\max_{x\in\bar{J}_i'}(P_n^{(d_l)}(x))^2 \le n\int_{x\in\mathbf{T}}(P_n^{(d_l)}(x))^2dx \le n\times n^{2d_l}\int_{x\in\mathbf{T}}(P_n(x))^2dx \le 2n^{2d_l+1}.$$

Hence by averaging we see that for least $k_l/2$ intervals J_i' in G_l satisfy

$$\max_{x\in J_i'}|P^{(d_l)}(x)| \le 2(n/k_l)^{1/2}n^{d_l}.$$

It follows from (11) and (12) that over these intervals

$$\max_{x\in J_i'}|P_n(x)| \le \frac{\lambda}{(n/2k_l)^{1/2}}(\frac{1}{n})^{d_l}2(n/k_l)^{1/2}n^{d_l} \le 4\lambda$$

and similarly,

$$\max_{x\in J_i'}|P_n'(x)| \le n\times\frac{\lambda}{(n/2k_l)^{1/2}}(\frac{1}{n})^{d_l}2(n/k_l)^{1/2}n^{d_l} \le 4\lambda n.$$

Letting A_l denote the union of all such intervals J_i' of a given good index l, and letting A denote the union of the A_l's over all good indices l, we obtain

$$\begin{aligned}\mathrm{Leb}(A) \ge \sum_{l,\mathrm{good}}(cd_l/n)k_l/2 \ge \sum_{l,\mathrm{good}}(c/4)d_{l+1}k_l/n \ge \sum_{l,\mathrm{good}}(c/4)m_l/n \\ \ge (c/4)(\varepsilon n/8 - \varepsilon n/32)/n \ge \frac{\varepsilon}{1024e}.\end{aligned}$$

Finally, notice that by definition of A we have

$$\max_{x\in A}|P_n(x)| \le 4\lambda \le \alpha \text{ and } \max_{x\in A}|P_n'(x)| \le 4\lambda n \le \beta n,$$

concluding the proof. □

We conclude the section by a quick consequence of Lemma 5.1. For each P_n that is not exceptional we let $N_s = N_s(P_n)$ and $N_{us} = N_{us}(P_n)$ be the number of roots of P_n over the set of stable and unstable intervals respectively.

Corollary 5.3 (Roots over stable intervals)**.** *With the same parameters as in Corollary 5.2, we have*

$$\mathbf{P}\Big(N_s 1_{P_n\in\mathcal{E}_e^c} \le \mathbf{E}N_n - \varepsilon n\Big) = o(1)$$

and

$$\mathbf{E}\Big(N_s 1_{P_n\in\mathcal{E}_e^c}\Big) \ge \mathbf{E}N_n - 2\varepsilon n/3.$$

Proof. For the first bound, by Corollary 5.2, if $N_s 1_{P_n\in\mathcal{E}_e^c} \le \mathbf{E}N_n - \varepsilon n$ then $N_n 1_{P_n\in\mathcal{E}_e^c} \le \mathbf{E}N_n - \varepsilon n/2$. Thus

$$\begin{aligned}\mathbf{P}\big(N_s 1_{P_n\in\mathcal{E}_e^c} \le \mathbf{E}N_n - \varepsilon n\big) &\le \mathbf{P}\big(N_n 1_{P_n\in\mathcal{E}_e^c} \le \mathbf{E}N_n - \varepsilon n/2\big)\\ &\le \mathbf{P}\big(\mathcal{E}_e^c \wedge N_n \le \mathbf{E}N_n - \varepsilon n/2\big) + \mathbf{P}(\mathcal{E}_e) = o(1),\end{aligned}$$

where we used (2) and Theorem 4.2.

For the second bound regarding $\mathbf{E}(N_s 1_{P_n\in\mathcal{E}_e^c})$, by Corollary 5.2, for non-exceptional P_n we have that $N_{us} \le \varepsilon n/2$, and hence trivially $\mathbf{E}(N_{us} 1_{P_n\in\mathcal{E}_e^c}) \le \varepsilon n/2$. Because each P_n has $O(n)$ roots, we then obtain

$$\begin{aligned}\mathbf{E}(N_s 1_{P_n\in\mathcal{E}_e^c}) &\ge \mathbf{E}N_n - \mathbf{E}(N_{us} 1_{P_n\in\mathcal{E}_e^c}) - \mathbf{E}(N_n 1_{P_n\in\mathcal{E}_e})\\ &\ge \mathbf{E}N_n - \varepsilon n/2 - O(n \times e^{-c\tau^4 n}) \ge \mathbf{E}N_n - 2\varepsilon n/3.\end{aligned}$$

□

6. Proof of the main results

We first give a deterministic result (see also [21, Claim 4.2]) to control the number of roots under perturbation.

Lemma 6.1. *Fix strictly positive numbers μ and ν. Let $I = (a, b)$ be an interval of length greater than $2\mu/\nu$, and let f be a C^1-function on I such that at each point $x \in I$ we have either $|f(x)| > \mu$ or $|f'(x)| > \nu$. Then for each root $x_i \in I$ with $x_i - a > \mu/\nu$ and $b - x_i > \mu/\nu$ there exists an interval $I(x_i) = (a', b')$ where $f(a')f(b') < 0$ and $|f(a')| = |f(b')| = \mu$, such that $x_i \in I(x_i) \subset (x_i - \mu/\nu, x_i + \mu/\nu)$ and the intervals $I(x_i)$ over the roots are disjoint.*

Proof. We may and will assume that f is not constant on I. By changing $f(x)$ to $\lambda_1 f(\lambda_2 x)$ for appropriate λ_1, λ_2, it suffices to consider $\mu = \nu = 1$. For each root x_i, and for $0 < t \le 1$ consider the interval $I_t(x_i)$ containing x_i of those points x where $|f(x)| < t$. We first show that for any $0 < t_1, t_2 \le 1$ and distinct roots $x_1, x_2 \in I$ satisfying the assumption of the lemma, we have that $I_{t_1}(x_1) \cap I_{t_2}(x_2) = \emptyset$. Assume otherwise, then because $f(x_1) = f(x_2) = 0$, there exists $x_1 < x < x_2$ such that $f'(x) = 0$ and $|f(x)| \le \min\{t_1, t_2\}$, which contradicts the assumption of the lemma. We will also show that $I_1(x_i) \subset (x_i - 1, x_i + 1)$. Indeed, assume otherwise for instance that $x_i - 1 \in I_1(x_i)$, then for all $x_i - 1 < x < x_i$ we have $|f(x)| < 1$, and so $|f'(x)| > 1$ over this interval. Without loss of generality we assume $f'(x) > 1$ for all x over this interval. The mean value theorem would then imply that $|f(x_i - 1)| = |f(x_i - 1) - f(x_i)| > 1$, a contradiction with $x_i - 1 \in I_1(x_i)$. As a consequence, we can define $I(x_i) = I_1(x_i)$, for which at the endpoints the function behaves as desired. □

Corollary 6.2. *Fix positive μ and ν. Let $I = (a, b)$ be an interval of length at least $2\mu/\nu$, and let f be a C^1-function on I such that at each point $x \in I$ we have either $|f(x)| > \mu$ or $|f'(x)| > \nu$. Let g be a function such that $|g(x)| < \mu$ over I. Then for each root $x_i \in I$ of f with $x_i - a > \mu/\nu$ and $b - x_i > \mu/\nu$ we can find a root x_i' of $f + g$ such that $x_i' \in (x_i - \mu/\nu, x_i + \mu/\nu)$, and also the x_i' are distinct.*

Proof. For each root $x_i \in I$ of f such that $x_i - a > \mu/\nu$ and $b - x_i > \mu/\nu$ we consider the interval $I(x_i) = (a', b')$ defined in Lemma 6.1. Without loss of generality, assume that $f(a') = -\mu$ and $f(b') = \mu$. Then as $|g(x)| < \mu$ over I, we have $f(a') + g(a') < 0 < f(b') + g(b')$. Hence by the mean value theorem there exists $x_i' \in (a', b')$ such that $f(x_i') + g(x_i') = 0$. We note that $I(x_i) \subset (x_i - \mu/\nu, x_i + \mu/\nu)$. □

Now we prove Theorem 1.4 by considering the two tails separately.

6.3. **The lower tail.** We need to show that for some constant c'

$$\mathbf{P}(N_n \le \mathbf{E}N_n - \varepsilon n) = O(e^{-c'\varepsilon^9 n}) \tag{13}$$

Here (and in (16)) c' depends on C_0. For instance it suffices to assume c' to be of order $1/C_0^4$ in the case of Theorem 2.5, and of order $1/C_0$ in the case of Theorem 2.6.

With the parameters $\alpha, \beta, \delta, \tau, R$ chosen as in Corollary 5.2, our key deterministic observation is that the number of roots in stable intervals of non-exceptional polynomials is not significantly decreased by small perturbations, as stated in the following.

Claim 6.4. *Let P_n be a non-exceptional polynomial such that P_n has at least $\mathbf{E}N_n - 2\varepsilon n/3$ roots over the stable intervals. Then for any trigonometric polynomial g of degree n with $\|g\|_2 \le \tau$, the polynomial $P_n + g$ has at least $\mathbf{E}N_n - \varepsilon n$ roots over $\mathbf{T}$.*

Proof. Let g be a trigonometric polynomial of degree n such that $\|g\|_2 \le \tau$. We first notice that the number of stable intervals I_j over which $\max_{x \in 3I_j} |g(x)| > \alpha$ is at most $O(\delta n)$. Indeed, assume that there are M such intervals I_j of length R/n. Then we can choose at least, say $M/6$, such intervals so that $3I_j$ are R/n-separated. By Theorem 2.2 we have

$$(M/6)\alpha^2 < n\tau^2,$$

which implies $M < 6n(\tau\alpha^{-1})^2 = O(\delta n)$.

As P_n is non-exceptional, there are at least $(\frac{2\pi}{R} - \delta)n$ stable intervals I_j. We will focus on the set $\mathcal{S}_{g,\alpha}$ of stable intervals I_j on which $\max_{x \in 3I_j} |g(x)| \le \alpha$, for which we have learned from the above argument that

$$|\mathcal{S}_{g,\alpha}| = (\frac{2\pi}{R} - O(\delta))n.$$

Notice furthermore that because P_n is non-exceptional and P_n has at least $\mathbf{E}N_n - 2\varepsilon n/3$ roots over the stable intervals, by Corollary 5.2, with appropriate choice of the parameters, P_n has at least $\mathbf{E}N_n - \varepsilon n$ roots over the stable intervals $I_j \in \mathcal{S}_{g,\alpha}$.

Continuing with the proof of the claim, for each $I_j \in \mathcal{S}_{g,\alpha}$, by Corollary 6.2 (applied to $I = 3I_j$ with $\mu = \alpha$ and $\nu = \beta n$, note that $\alpha/\beta \asymp \delta^{3/4} < R$ and $\max_{x \in 3I_j} |g(x)| < \alpha$), the number of roots of $P_n + g$ over I_j is at least as that of P_n. As a consequence, as P_n has at least $\mathbf{E}N_n - \varepsilon n$ roots over the stable intervals $I_j \in \mathcal{S}_{g,\alpha}$, the perturbed polynomial $P_n + g$ also has at least $\mathbf{E}N_n - \varepsilon n$ roots over these stable intervals I_j of $\mathcal{S}_{g,\alpha}$. In particularly $P_n + g$ has at least $\mathbf{E}N_n - \varepsilon n$ roots over $\mathbf{T}$. □

We next turn to the probabilistic part of the argument, uses the above stability over perturbations. Let $\mathcal{U}^{lower}$ be the collection of $\mathbf{v}_{P_n}$ in Ω^{2n} from non-exceptional P_n that have at least $\mathbf{E}N_n - 2\varepsilon n/3$ roots over the stable intervals. Then by Corollary 5.3

$$\mathbf{P}(\mathbf{v}_{P_n} \in \mathcal{U}^{lower}) \ge 1 - \mathbf{P}\big(\mathcal{E}_e\big) - \mathbf{P}\big(N_s 1_{P_n \in \mathcal{E}_e^c} \le \mathbf{E}N_n - 2\varepsilon n/3\big) \ge 1/2. \tag{14}$$

We can now complete the proof of (13). By Claim 6.4 the set $\{\mathbf{v} \in \Omega^{2n}, d_2(\mathbf{v}, \mathcal{U}^{lower}) \le \tau\sqrt{2n}\}$ is contained in the set of having at least $\mathbf{E}N_n - \varepsilon n$ roots. Furthermore, (14) says that $\mathbf{P}(\mathbf{v}_{P_n} \in$

$\mathcal{U}^{lower}) \geq 1/2$. Hence by Theorems 2.5 and 2.6

$$\mathbf{P}(N_n \geq \mathbf{E}N_n - \varepsilon n) \geq \mathbf{P}\Big(\mathbf{v}_{P_n} \in \big\{\mathbf{v} \in \Omega^{2n}, d_2(\mathbf{v}, \mathcal{U}^{lower}) \leq \tau\sqrt{2n}\big\}\Big) \geq 1 - 2\exp(-c'\varepsilon^9 n), \tag{15}$$

where we used the fact that $\tau \asymp \delta^2$ from (9) and that δ satisfies (10). Here we can take c' to be of order $1/C_0^4$ in the case of Theorem 2.5, and of order $1/C_0$ in the case of Theorem 2.6. □

6.5. **The upper tail.** Our goal here is to justify the upper tail

$$\mathbf{P}(N_n \geq \mathbf{E}N_n + \varepsilon n) = O(e^{-c'\varepsilon^9 n}). \tag{16}$$

Let $\mathcal{U}^{upper}$ denote the set of $\mathbf{v}_{P_n}$ in Ω^{2n} for which $N_n \geq \mathbf{E}N_n + \varepsilon n$ and P_n is non-exceptional. By Theorem 4.2, for (16) it suffices to show that

$$\mathbf{P}(\mathbf{v}_{P_n} \in \mathcal{U}^{upper}) = O(e^{-c'\varepsilon^9 n}). \tag{17}$$

Proof. (of Equation (17)) Assume that for a non-exceptional P_n we have $N_n \geq \mathbf{E}N_n + \varepsilon n$. Then by Lemma 5.1 (Corollary 5.2) the number of roots of P_n over the stable intervals is at least $\mathbf{E}N_n + 2\varepsilon n/3$. Arguing as in the previous subsection (with the same parameters of $\alpha, \beta, \tau, \delta$), Corollary 5.2 and Corollary 6.2 imply that $h = P_n + g$ with any g satisfying $\|g\|_2 \leq \tau$ has at least $\mathbf{E}N_n + \varepsilon n/2$ roots. On the other hand, we know by (2) that the probability that P_n belongs to this set of trigonometric polynomials is smaller than $1/2$. It thus follows by Theorems 2.5 and 2.6 that

$$\mathbf{P}(\mathbf{v}_{P_n} \in \mathcal{U}^{upper}) = O(e^{-c'\varepsilon^9 n}), \tag{18}$$

where we again used that $\tau \asymp \delta^2$ and δ satisfies (10), and where c' depends on C_0 as in (15). □

We conclude this section with some comments regarding Remark 1.5. Note that Theorem 3.1, and all other deterministic results still work under this setting. Choose $C_0 = n^{c_0}$ where the positive constant c_0 is chosen to be small enough (depending on the exponent of the sub-exponential tail, or on C). It then suffices to establish the result for $|\xi| \leq n^{c_0}$. We next choose the parameters α, β, δ as in Theorem 4.2, where say $\delta = n^{-c_0}$. With these parameters, we note that Theorem 4.2 still holds with sub-exponential decay of rate $O(\exp(-n^{1-12c_0}))$, and Equations (13) and (16) still hold with sub-exponential decay of rate $O(\exp(-n^{1-4c_0}\varepsilon^9))$.

Appendix A. Concentration results

Proof. (of Theorem 2.5) Consider the function $F(\boldsymbol{\xi}) := d_1(\boldsymbol{\xi}, \mathcal{A})$, which measures the L_1-distance. This function is $2C_0$-Lipschitz (coordinatewise), so by McDiarmid's inequality, with $\mu = \mathbf{E}F(\boldsymbol{\xi})$

$$\mathbf{P}(|F(\boldsymbol{\xi}) - \mu| \geq \lambda) \leq 2\exp(-\lambda^2/2nC_0^2).$$

This then implies that

$$\mathbf{P}(F(\boldsymbol{\xi}) = 0)\mathbf{P}(F(\boldsymbol{\xi}) \geq \lambda) \leq 4\exp(-\lambda^2/4nC_0^2).$$

Indeed, if $\lambda \leq \mu$ then

$$\mathbf{P}(F(\boldsymbol{\xi}) = 0) \leq \mathbf{P}(F(\boldsymbol{\xi}) - \mu \leq -\mu) \leq 2\exp(-\mu^2/2nC_0^2) \leq 2\exp(-\lambda^2/2nC_0^2),$$

while if $\lambda \geq \mu$ then

$$\begin{aligned}\mathbf{P}(F(\boldsymbol{\xi}) = 0)\mathbf{P}(F(\boldsymbol{\xi}) \geq \lambda) &\leq \mathbf{P}(F(\boldsymbol{\xi}) - \mu \leq -\mu)\mathbf{P}(F(\boldsymbol{\xi}) - \mu \geq \lambda - \mu) \\ &\leq 4\exp(-(\mu^2 + (\lambda-\mu)^2)/2nC_0^2) \leq 4\exp(-\lambda^2/4nC_0^2).\end{aligned}$$

Now because of boundedness (where we recall that $\mathcal{A} \subset \Omega^n = [-C_0, C_0]^n$), for any $\mathbf{y} \in \mathcal{A}$,

$$\|\boldsymbol{\xi} - \boldsymbol{\xi}'\|_2^2 \le 2C_0 \|\boldsymbol{\xi} - \boldsymbol{\xi}'\|_1.$$

So if $d_2(\boldsymbol{\xi}, \mathcal{A}) \ge t\sqrt{n}$ then $d_1(\boldsymbol{\xi}, \mathcal{A}) \ge t^2 n/2C_0$. We thus obtain

$$\mathbf{P}(\boldsymbol{\xi} \in \mathcal{A})\mathbf{P}(d_2(\boldsymbol{\xi}, \mathcal{A}) \ge t\sqrt{n}) \le \mathbf{P}(\boldsymbol{\xi} \in \mathcal{A})\mathbf{P}(d_1(\boldsymbol{\xi}, \mathcal{A}) \ge t^2 n/2C_0) \le 4\exp(-t^4 n/16C_0^4).$$

□

Proof. (of Theorem 2.6) Let $\lambda := t\sqrt{n}$ and $F(\boldsymbol{\xi}) := \min\{d_2(\boldsymbol{\xi}, \mathcal{A}), \lambda\}$. Then F is 1-Lipschitz, and

$$\mathbf{E}F(\boldsymbol{\xi}) \le (1 - \mathbf{P}(\boldsymbol{\xi} \in \mathcal{A}))\lambda.$$

It is known (see for instance [19]) that for distributions satisfying log-Sobolev inequality we have that

$$\mathbf{P}(F(\boldsymbol{\xi}) \ge \mathbf{E}F(\boldsymbol{\xi}) + t) \le \exp(-t^2/4C_0).$$

Thus, since $\mathbf{E}F(\boldsymbol{\xi}) = \mathbf{P}(\boldsymbol{\xi} \notin \mathcal{A})\mathbf{E}(F(\boldsymbol{\xi})|\boldsymbol{\xi} \notin \mathcal{A}) \le \lambda \mathbf{P}(\boldsymbol{\xi} \notin \mathcal{A})$,

$$\begin{aligned}\mathbf{P}(d_2(\boldsymbol{\xi}, A) \ge \lambda) = \mathbf{P}(F(\boldsymbol{\xi}) \ge \lambda) &\le \mathbf{P}(F(\boldsymbol{\xi}) \ge \mathbf{E}F(\mathbf{x}) + \mathbf{P}(\boldsymbol{\xi} \in \mathcal{A})\lambda) \\ &\le \exp(-\mathbf{P}^2(\boldsymbol{\xi} \in \mathcal{A})\lambda^2/4C_0).\end{aligned}$$

Finally, to see the last claim in Theorem 2.6 we let $\mathcal{A}' = \{\mathbf{a}, d_2(\mathbf{a}, \mathcal{A}) \ge t\sqrt{n}\}$. By the theorem's main conclusion, because $\mathbf{P}(\boldsymbol{\xi} \in \mathcal{A}') \ge 1/2$, we have

$$\mathbf{P}(d_2(\boldsymbol{\xi}, \mathcal{A}') \ge t\sqrt{n}) \le 2\exp(-t^2 n/16C_0).$$

To this end, we just observe that if $\mathbf{a} \in \mathcal{A}$ then $d_2(\mathbf{a}, \mathcal{A}') \ge t\sqrt{n}$ (otherwise there would exist $\mathbf{a}' \in \mathcal{A}'$ such that $d_2(\mathbf{a}, \mathbf{a}') < t\sqrt{n}$, which would imply $d_2(\mathbf{a}', \mathcal{A}) < t\sqrt{n}$, a contradiction with the definition of $\mathcal{A}'$.) □

Acknowledgements. The authors are grateful to O. Nguyen and T. Erdély for help with references. They also thank the anonymous referees for helpful suggestions.

References

[1] J. Angst and G. Poly, A weak Cramér condition and application to Edgeworth expansions, Electron. J. Probab. Volume 22 (2017), paper no. 59, 24 pp.

[2] J. Angst and G. Poly, Variations on Salem–Zygmund results for random trigonometric polynomials: application to almost sure nodal asymptotics, Electron. J. Probab. 26: 1-36 (2021).

[3] J. M. Azaïs, F. Dalmao and J. León, CLT for the zeros of classical random trigonometric polynomials. Ann. Inst. Henri-Poincare. 52(2) (2016), 804-820.

[4] J. M. Azaïs and J. León, CLT for crossings of random trigonometric polynomials. Electron. J. Probab. 18 (68) (2013), 1-17.

[5] R. Basu, A. Dembo, N. Feldheim and O. Zeitouni, Exponential concentration for zeroes of stationary Gaussian processes, Int. Math. Res. Notices 23 (2020), 9769-9796.

[6] R. N. Bhattacharya and R. Rao. Normal approximation and asymptotic expansions. Robert E. Krieger Publishing Co., Inc., Melbourne, FL, 1986.

[7] V. Bally, L. Caramellino and G. Poly, Non universality for the variance of the number of real roots of random trigonometric polynomials, Probab. Theory Relat. Fields 174 (2019), 887–927.

[8] P. Borwein and T. Erdélyi, Polynomials and Polynomial Inequalities, Graduate texts in Mathematics, Springer (1995), Berlin–New York.

[9] Y. Do, H. Nguyen and V. Vu. Real roots of random polynomials: expectation and repulsion, Proceedings London Mathematical Society (2015), Vol. 111 (6), 1231-1260.

[10] Y. Do, H. Nguyen and O. Nguyen, Random trigonometric polynomials: universality and non-universality of the variance of the number of real roots, Annales de l'Institut Henri Poincare, Vol. 58, No. 3 (2022), 1460–1504.

[11] 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.
[12] C. G. Esseén, On the Kolmogorov-Rogozin inequality for the concentration function, Z. Wahrsch. Verw. Gebiete 5 (1966), 210-216.
[13] O. N. Feldheim and A. Sen, Double roots of random polynomials with integer coefficients, Electron. J. Probab. Volume 22 (2017), paper no. 10, 23 pp.
[14] D. Gayet and J.-Y. Welschinger, Exponential rarefaction of real curves with many components, Publ. Math. IHES 113 (2011), 69-93.
[15] A. Granville and I. Wigman. The distribution of the zeros of random trigonometric polynomials. Amer. J. Math. 133 (2) (2011) 295-357.
[16] H. Iwaniec and E. Kowalski, Analytic number theorem, Colloqium Publications 53, AMS (2004), Providence, RI.
[17] G. Halász, Estimates for the concentration function of combinatorial number theory and probability, Period. Math. Hungar. 8 (1977), no. 3-4, 197-211.
[18] S. V. Konyagin and W. Schlag, Lower bounds for the absolute value of random polynomials on a neighborhood of the unit circle, Transactions AMS 351 (1999), 4963-4980.
[19] M. Ledoux, The concentration of measure phenomenon, Mathematical Surveys and Monographs 89, AMS (2001), Providence, RI.
[20] D. S. Lubinsky, A. Mate and P. Nevai, Quadrature sums involving p-th powers of polynomials, SIAM J. Math. Anal. 18(1987) 531-544.
[21] F. Nazarov and M. Sodin, On the number of nodal domains of random spherical harmonics. Amer. J. Math. 131 (2009), 1337-1357.
[22] F. Nazarov and M. Sodin, Fluctuations in random complex zeroes: asymptotic normality revisited, Int. Math. Res. Notices 24 (2011), 5720-5759.
[23] H. Nguyen, O. Nguyen and V. Vu, On the number of real roots of random polynomials, Communications in Contemporary Mathematics (2016) Vol. 18, 4, 1550052.
[24] O. Nguyen and V. Vu, Roots of random functions: a framework for local universality, American Journal of Mathematics 144(01), 2022, 1-74.
[25] H. H. Nguyen and O. Zeitouni, Exponential concentration for the number of roots of random trigonometric polynomials. arXiv:1912.12051v1.
[26] R. Peled, A. Sen and O. Zeitouni, Double roots of random Littlewood polynomials, Israel Journal of Mathematics, (213) 2016, 55-77.
[27] C. Qualls, On the number of zeros of a stationary Gaussian random trigonometric polynomial, J. London Math. Soc. (2) 2 (1970), 216-220.
[28] G. Szegő, Orthogonal Polynomial, 4th ed. American Mathematical Society (1975), Providence, RI.
[29] T. Tao and V. Vu, Random matrices: The Circular Law, Communication in Contemporary Mathematics 10 (2008), 261-307.
[30] T. Tao and V. Vu, Local universality of zeroes of random polynomials. Int. Math. Res. Notices, paper rnu084, 2014.
[31] P. Tchebycheff. Sur deux théoremes relatifs aux probabilités. Acta Math., 14(1) (1890), 305-315.

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

Email address: nguyen.1261@osu.edu

Faculty of Mathematics, Weizmann Institute and Courant Institute, NYU, Rehovot 76100, Israel and NY 10012, USA

Email address: ofer.zeitouni@weizmann.ac.il