Random polynomials having few or no real zeros

Amir Dembo * Bjorn Poonen † Qi-Man Shao ‡ Ofer Zeitouni §

May 29, 2000

Abstract

Consider a polynomial of large degree n whose coefficients are independent, identically distributed, nondegenerate random variables having zero mean and finite moments of all orders. We show that such a polynomial has exactly k real zeros with probability $n^{-b+o(1)}$ as $n \to \infty$ through integers of the same parity as the fixed integer $k \geq 0$. In particular, the probability that a random polynomial of large even degree n has no real zeros is $n^{-b+o(1)}$. The finite, positive constant b is characterized via the centered, stationary Gaussian process of correlation function $\operatorname{sech}(t/2)$. The value of b depends neither on k nor upon the specific law of the coefficients. Under an extra smoothness assumption about the law of the coefficients, with probability $n^{-b+o(1)}$ one may specify also the approximate locations of the k zeros on the real line. The constant b is replaced by $b/2$ in case the i.i.d. coefficients have a nonzero mean.

1 Introduction

Let $\{a_i\}_{i=0}^{\infty}$ denote a sequence of independent, identically distributed (i.i.d.) random variables of zero mean and unit variance. Consider the random polynomial

$$f_n(x) = \sum_{i=0}^{n-1} a_i x^i . \tag{1.1}$$

For n odd, define

$$P_n = \mathrm{P}\left(f_n(x) > 0 \quad \forall x \in \mathbb{R}\right). \tag{1.2}$$

As described in Section 1.1, the study of the number of roots of random polynomials has a long history. Our main goal is to prove that $P_n = n^{-b+o(1)}$ as $n \to \infty$ for a finite constant $b > 0$, at least when the coefficient distribution has finite moments of all orders. The constant b can be described in terms of the centered stationary Gaussian process Y_t with correlation function $R_y(t) = \operatorname{sech}(t/2)$ (see (1.4) for an explicit construction of $Y_.$). Define

$$b = -4 \lim_{T\to\infty} \frac{1}{T} \log P\Big(\sup_{0\leq t\leq T} Y_t \leq 0\Big) , \tag{1.3}$$

---

*Research partially supported by NSF grant DMS-9704552.

†Supported by NSF grant DMS-9801104, a Sloan Fellowship, and a Packard Fellowship.

‡Research partially supported by NSF grant DMS-9802451.

§Research partially supported by a grant from the Israel Science Foundation and by the fund for promotion of research at the Technion.

AMS subject classification: primary 60G99; secondary 12D10, 26C10.
Keywords: Random polynomials, Gaussian processes

where, throughout this paper, log denotes the natural logarithm. The existence of the limit in (1.3) and the estimate $b \in [0.4, 2]$ are proved in Lemma 2.5. We note in passing that our numerical simulations of random polynomials of degree $n-1 \leq 1024$ suggest $b = 0.76 \pm 0.03$.

Our main result, which is a consequence of Theorem 1.3 stated in Section 1.2, is the following

Theorem 1.1 *a) Suppose $\{a_i\}$ is a sequence of zero-mean, unit-variance, i.i.d. random variables possessing finite moments of all order. Then,*

$$\lim_{n\to\infty} \frac{\log P_{2n+1}}{\log n} = -b\,.$$

b) If $\{a_i\}$ is as above but with $E(a_i) = \mu \neq 0$, we denote $P_n^\mu = \mathrm{P}\left(f_n(x) \neq 0 \quad \forall x \in \mathbb{R}\right)$. Then,

$$\lim_{n\to\infty} \frac{\log P_{2n+1}^\mu}{\log n} = -b/2\,.$$

It is interesting to note that one may answer questions related to a prescribed number of zeros. Our main result in this direction is the following theorem. For a slightly different variant, allowing to prescribe the *location* of zeros, see also Proposition 1.5.

Theorem 1.2 *Under the assumptions of Theorem 1.1 a), the probability that the random polynomial $f_{n+1}(x)$ of degree n has $o(\log n/\log\log n)$ real zeros is $n^{-b+o(1)}$ as $n \to \infty$. For any fixed k, the probability $p_{n,k}$ that f_{n+1} has exactly k real zeros, all of which are simple, satisfies*

$$\lim_{n\to\infty} \frac{\log p_{2n+k,k}}{\log n} = -b\,.$$

(Obviously, $p_{n,k} = 0$ when $n-k$ is odd.)

The key to our analysis is a detailed study of the case where the coefficients are Gaussian, implying that $f_n(\cdot)$ is a Gaussian process (Gaussian processes are particularly useful in this context because for them comparisons can be made via Slepian's lemma). The extension to general distribution uses the strong aproximation results of Komlós-Major-Tusnády [KMT]. Although this technique requires finite moments of all order, we conjecture that the asymptotic $n^{-b+o(1)}$ applies to $p_{n,k}$ for $n-k$ even, whenever the nondegenerate zero-mean i.i.d. a_i are in the domain of attraction of the Normal distribution. This conjecture is supported by the following heuristic derivation of $P_n = n^{-b+o(1)}$.

For $x \in [0,1]$ near 1, let $x = 1 - e^{-t}$. Note that $x^i \approx \exp(-e^{-t} i)$ when $t \gg 0$, and moreover, the function $h_t(u) := \exp(-e^{-t} u)$ changes slowly in u for $t \gg 0$. Summation by parts suggests that the sign of $f_n(x)$ is mostly determined by the behavior of $\sum_{i=0}^{j} a_i$ for large j depending on t. Hence, for a_i in the domain of attraction of the Normal distribution, we next replace a_i with i.i.d. standard Normal variables b_i. Using the representation $b_i = W_{i+1} - W_i$ for a standard Brownian motion W_t we further replace the sum over $i = 0, \ldots, n-1$ with the corresponding stochastic integral over $[0, \infty)$. This in turn yields the approximation of the normalized $\hat{f}_n(x) := f_n(x)/\sqrt{\mathrm{Var}(f_n(x))}$ by the centered, Gaussian process

$$Y_t = \frac{\int_0^\infty h_t(u) dW_u}{(\int_0^\infty h_t(u)^2 du)^{1/2}}\,. \tag{1.4}$$

It is easy to check that the process $Y_\cdot$ of (1.4) is stationary, with correlation function $\mathrm{sech}(t/2)$. By continuity arguments, $f_n(x)$ typically has a constant sign in $[1 - n^{-1}, 1]$, so our approximation procedure is relevant only as long as $t \leq \log n$. Alternatively, $t = \log n$ is where we start seeing $h_t(n) = O(1)$, contrasting the replacement of the upper limit n in the discrete sum with the upper limit ∞ in the stochastic integral of (1.4). We are to consider the possibility of $f_n(x) = 0$ for x in the left and in the right neighborhoods of both -1 and $+1$. In each of these four regimes of x the function $\hat{f}_n(x)$ is amenable to a similar treatment, leading to an approximation by the process Y_t. With $\hat{f}_n$ having approximately independent values in the four different regimes, we arrive at the formula (1.3) for b.

It is natural to wonder what happens when a_i are of a symmetric law that is in the domain of attraction of an α-stable law, for some $\alpha \in (0, 2)$. A lower bound on P_n of the form n^{-c} for some finite value of c is then easily obtained by considering the event that a_0 and a_{n-1} are "huge" and positive, while other coefficients are "reasonable." Repeating the above heuristic for this case, one is led to believe that the formula (1.3) still applies, but now with Y_t of (1.4) replaced by

$$Y_{t,\alpha} = \frac{\int_0^\infty h_t(u) dX_u^{(\alpha)}}{(\int_0^\infty h_t(u)^\alpha du)^{1/\alpha}} \,, \tag{1.5}$$

where $X^{(\alpha)}_\cdot$ denotes the symmetric stable process of index α and the stochastic integral in (1.5) is to be interpreted via integration by parts. We have yet no strong evidence to support the above statement. However, our numerical simulations indicate the behavior $P_n = n^{-b_1+o(1)}$ for i.i.d. Cauchy random variables a_i (that is, $\alpha = 1$), where $b_1 \approx 0.86$ is larger than b.

1.1 Historical remarks

Throughout this section, $\{a_i\}$ are independent, identically distributed, nondegenerate, real-valued random variables.[1] Let N_n denote the number of distinct[2] real zeros of $f(x) := \sum_{i=0}^n a_i x^i$. (For the sake of definiteness, we define $N_n = 0$ when f is the zero polynomial.) So, $p_n := P(N_n = 0) = p_{n,0}$ and we also let E_n and V_n denote the mean and variance of N_n.

The study of real roots of random polynomials has a long and full history, but most of it deals with the asymptotic behavior of E_n instead of p_n. Presumably this is because E_n is much easier to estimate: because expectation is linear, one can compute E_n by integrating over the real line the probability of having a root in $(t, t + dt)$, for example.

Although as mentioned in [To, p. 618], one can find probabilistic statements in the context of roots of polynomials as early as 1782 (Waring) and 1864 (Sylvester), the first people to study the asymptotic behavior of N_n seem to be Bloch and Pólya [BP]. In 1932, they proved $E_n = O(n^{1/2})$ for the coefficient distribution $P(a_i = 1) = P(a_i = 0) = P(a_i = -1) = 1/3$. This work led Littlewood and Offord to undertake a systematic study of N_n in a series of papers [LO1],[LO2],[LO3] starting in 1938. They proved that if the a_i are all uniform on $[-1, 1]$, or all Normal, or all uniform on $\{-1, 1\}$, (i.e. $P(a_i = 1) = P(a_i = -1) = 1/2$), then

$$P\left(N_n > 25(\log n)^2\right) \le \frac{12 \log n}{n}, \quad \text{and} \quad P\left(N_n < \frac{\alpha \log n}{(\log \log n)^2}\right) < \frac{A}{\log n}$$

for some constants α and A. In particular, for some constant α',

$$\frac{\alpha' \log n}{(\log \log n)^2} \le E_n \le 25(\log n)^2 + 12 \log n$$

[1] Some authors whose work we mention assumed $a_0 = 1$ or $a_n = 1$, but as far as asymptotic behavior as $n \to \infty$ is concerned, it makes little difference.

[2] The asymptotic behavior does not depend on whether roots are counted with multiplicity or not.

and $p_n = O(1/\log n)$ for these distributions. This upper bound for p_n has apparently not been improved, until the current paper.[3]

In 1943 Kac [Ka1] found the exact formula

$$E_n = \frac{1}{\pi}\int_{-\infty}^{\infty}\sqrt{\frac{1}{(t^2-1)^2} - \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}}dt\ , \tag{1.6}$$

when a_i is Normal with mean zero, and extracted from it the asymptotic estimate

$$E_n \sim \frac{2}{\pi}\log n\ . \tag{1.7}$$

Much later Jamrom [Ja] and Wang [Wa] improved this to $E_n = (2/\pi)\log n + C + o(1)$ for an explicit constant C, and ultimately Wilkins [Wi] obtained an asymptotic series for E_n from (1.6). In 1949 Kac [Ka2] obtained (1.7) for the case where a_i is uniform on $[-1,1]$. Erdős and Offord [EO] obtained the same asymptotics for a_i uniform on $\{-1,1\}$. Stevens [St] proved (1.7) for a wide class of distributions, and this estimate was finally extended by Ibragimov and Maslova [IM1],[IM2] to all mean-zero distributions in the domain of attraction of the Normal law.

At around the same time (the late 1960's), Logan and Shepp [LS1],[LS2] discovered that if the coefficient distribution is the symmetric stable distribution with characteristic function $\exp(-|z|^\alpha)$, $0<\alpha\le 2$, then $E_n \sim c_\alpha \log n$, where

$$c_\alpha := \frac{4}{\pi^2\alpha^2}\int_{-\infty}^{\infty} dx\ \log\int_0^\infty \frac{|x-y|^\alpha e^{-y}}{|x-1|^\alpha}\,dy > \frac{2}{\pi}.$$

They also proved $\lim_{\alpha\to 0^+} c_\alpha = 1$, and performed calculations that suggested that c_α is a decreasing function of α, terminating at $c_2 = 2/\pi$, Kac's value for the Normal distribution. Ibragimov and Maslova [IM4] extended these results by finding the asymptotic behavior of E_n for arbitrary distributions in the domain of attraction of a stable distribution. The asymptotics are different when the distribution has nonzero mean; for instance [IM3], if a_i are Normal with nonzero mean, then $E_n \sim (1/\pi)\log n$ instead of $(2/\pi)\log n$. Shepp (private communication) has conjectured that there exists a universal constant B such that

$$\limsup_{n\to\infty}\frac{E_n}{\log n} \le B$$

for *any* coefficient distribution (satisfying only the hypotheses at the beginning of this section). If B exists, then $B \ge 1$ by the work of Logan and Shepp mentioned above.

In 1974, Maslova [Ma1],[Ma2] proved that if $P(a_i=0)=0$, $Ea_i = 0$ and $E(a_i^{2+\epsilon})<\infty$ for some $\epsilon>0$, then

$$V_n \sim \frac{4}{\pi}\left(1-\frac{2}{\pi}\right)\log n \tag{1.8}$$

and N_n is asymptotically Normal.

Much work was also done on complex roots of f_n; see [IZ] and references therein for an updated account. Further results on random polynomials and their generalizations can be found in the books [BR, Fa] and the survey article [EK].

Our interest in the asymptotics of p_n grew out of a problem in arithmetic geometry. The paper [PS] showed that Jacobians of curves over $\mathbb{Q}$ could be odd, in the sense of having Shafarevich-Tate groups of nonsquare order (despite prior claims in the literature that this was impossible).

[3] The only result in the literature that might be said to have improved our knowledge of p_n is (1.8), which together with (1.7) implies for many distributions that $\limsup_{n\to\infty} p_n \log n \le \pi - 2$. The bound has the same form as that arising from the work of Littlewood and Offord, but the constant has been made explicit.

Moreover it was shown (in a sense that was made precise) that the probability that a random hyperelliptic curve $y^2 = f(x)$ of genus g over $\mathbb{Q}$ has odd Jacobian could be related to a sequence of "local" probabilities, one for each nontrivial absolute value on $\mathbb{Q}$. The computation of the local probability for the standard archimedean absolute value reduced to the knowledge of the probability that the curve $y^2 = f(x)$ has no real point, or equivalently, the probability that the random polynomial $f(x)$ satisfies $f(x) < 0$ for all real x. Although the asymptotic behavior of this probability was not needed in a substantial way in [PS], the authors of that paper found the question to be of sufficient interest in its own right that they developed heuristics that led them to conjecture the existence of a universal constant $b > 0$ such that $p_n = n^{-b+o(1)}$, for any mean-zero distribution in the domain of attraction of the Normal law.

1.2 Statement of main theorems

Let $\hat{f}_n(x) := f_n(x)/\sqrt{E(f_n(x)^2)}$ denote the normalized random polynomial, so $\hat{f}_n(x)$ has unit variance for each x. Instead of proving only $P_n = n^{-b+o(1)}$, we generalize in the following, to facilitate applications to related problems.

Theorem 1.3 *Suppose a_i are zero-mean i.i.d. random variables of unit variance and with finite moments of all orders. For $n-1$ even, let*

$$P_{n,\gamma_n} = P\Big(\hat{f}_n(x) > \gamma_n(x) \quad \forall x \in \mathbb{R}\Big)\ ,$$

for nonrandom functions $\gamma_n(x)$ such that $n^\delta|\gamma_n(x)| \to 0$ uniformly in $x \in \mathbb{R}$ for some $\delta > 0$. Then,

$$\frac{\log P_{n,\gamma_n}}{\log n} \underset{n\to\infty}{\longrightarrow} -b\ . \tag{1.9}$$

The upper bound $P_{n,\gamma_n} \le n^{-b+o(1)}$ applies as soon as

$$\inf\{\gamma_n(x) : ||x|-1| \le n^{-\varepsilon_n}\} \to 0 \qquad \text{for any } \varepsilon_n \to 0\ .$$

The key to the proof of Theorem 1.3 is the analysis of P_{n,γ_n} for random polynomials $f_n(x)$ with coefficients $\{b_i\}$ that are i.i.d. standard Normal variables. To distinguish this case, we use throughout the notations $f_n^b(x)$, $\hat{f}_n^b$ and P_{n,γ_n}^b for $f_n(x)$, $\hat{f}_n(x)$ and P_{n,γ_n}, respectively, when dealing with polynomials of coefficients that are Normal variables. The next theorem summarizes our results in this special case.

Theorem 1.4 *The convergence of $\log P_{n,\gamma_n}^b/\log n$ to $-b$ applies in the standard Normal case, as soon as the nonrandom functions $\gamma_n(x) \le M < \infty$ are such that*

$$\sup\{|\gamma_n(x)| : ||x|-1| \le n^{-\varepsilon_n}\} \to 0 \qquad \text{for some } \varepsilon_n \to 0\ .$$

The following proposition is the variant of Theorem 1.2 alluded to above. It shows that with probability $n^{-b+o(1)}$ one may also prescribe arbitrarily the location of the k real zeros of $f_{n+1}(x)$, provided the support of the law of a_i contains an open interval around 0. The latter assumption is to some extent necessary. For example, when $P(a_i = 1) = P(a_i = -1) = 1/2$ it is easy to see that $f_{n+1}(x)$ cannot have zeros in $[-1/2, 1/2]$.

Proposition 1.5 *Suppose a_i are zero-mean i.i.d. random variables of unit variance, finite moments of all orders, and the support of the law of each a_i contains the interval $(-\eta,\eta)$ for some $\eta > 0$. Given disjoint open intervals $U_1, \ldots, U_\ell$ and positive integers $m_1, \ldots, m_\ell$, the probability that the random polynomial $f_{n+1}(x)$ has exactly m_i real zeros in U_i for each i and no real zeros anywhere else is $n^{-b+o(1)}$ for $n \to \infty$ through integers of the same parity as $k = \sum_i m_i$.*

The organization of this paper is as follows. Auxiliary lemmas about Gaussian processes, needed for the proof of Theorem 1.4, are grouped in Section 2 (including the bounds on b mentioned in the introduction, c.f. Lemma 2.5). Relying upon Gaussian techniques, the proof of the lower bound of Theorem 1.4 is in Section 3, and the complementary upper bound in Section 4. Building upon Theorem 1.4, and with the help of strong approximation, Section 5 provides the proof of our main result, Theorem 1.3. Theorem 1.1 is then derived in Section 6. Section 7 provides the upper bound on the probability of interest in Theorem 1.2, with the lower bound proved in Section 8. Finally, Proposition 1.5 is proved in Section 9.

2 Auxiliary lemmas

We start by introducing several notations that appear throughout this work. For n odd, let $c_n(x,y)$ denote the covariance function of $\hat{f}_n(x)$, that is

$$c_n(x,y) = \frac{E\Big(f_n(x)f_n(y)\Big)}{\sqrt{E(f_n(x)^2)E(f_n(y)^2)}} \tag{2.1}$$

Then, for $x \neq \pm 1$ and $y \neq \pm 1$,

$$c_n(x,y) = \frac{g(x^n,y^n)}{g(x,y)} \tag{2.2a}$$

where

$$g(x,y) = \frac{|xy-1|}{\sqrt{|(1-x^2)(1-y^2)|}} \geq 0\,. \tag{2.2b}$$

Note that $g(x,y) = g(-x,-y) = g(\frac{1}{x},\frac{1}{y})$. Further,

$$\forall x,y \in (-1,1), \quad g(x,y) \geq 1 \tag{2.2c}$$

and the change of variables $z = 1-x, w = 1-y$, leads to

$$\frac{1}{g(x,y)} = \frac{2\sqrt{zw}}{z+w}\Big[1-\Big[\frac{1-\frac{wz}{z+w}-\sqrt{1-\frac{z}{2}}\,\sqrt{1-\frac{w}{2}}}{1-\frac{wz}{z+w}}\Big]\Big] \tag{2.3}$$

A good control on $g(x,y)$ is provided by the following lemma.

Lemma 2.1 *For any $z, w \in (0,1/2]$*

$$\frac{1}{8}(w-z)^2 \leq \Big(1-\frac{wz}{w+z}-\sqrt{1-\frac{z}{2}}\,\sqrt{1-\frac{w}{2}}\Big)\frac{\max(z,w)}{\Big(1-\frac{wz}{w+z}\Big)} \leq (w-z)^2\,.$$

Proof: Let $z+w=\eta$, $z-w=\xi$, assuming without loss of generality that $0<\xi\leq\eta\leq 1$. Since

$$1 \geq 1-\frac{wz}{w+z} \geq \frac{1}{2}\,, \quad 1 \geq \frac{\max(z,w)}{(z+w)} \geq \frac{1}{2}\,,$$

it suffices to prove that

$$f(w,z) = \frac{1}{(z-w)^2}\Big[1-\frac{wz}{w+z}-\sqrt{1-\frac{z}{2}}\,\sqrt{1-\frac{w}{2}}\,\Big](z+w) \in [1/4,1/2]\,.$$

To this end, observe that for all $0 < \xi \le \eta \le 1$ we have

$$\begin{aligned} f(w,z) &= \frac{\eta}{\xi^2}\Big[1 - \frac{\eta^2-\xi^2}{4\eta} - \sqrt{\Big(1-\frac{\eta}{4}\Big)^2 - \Big(\frac{\xi}{4}\Big)^2}\,\Big] \\ &= \frac{\eta}{\xi^2}\Big[\Big(1-\frac{\eta}{4}\Big) + \frac{\xi^2}{4\eta} - \sqrt{\Big(1-\frac{\eta}{4}\Big)^2 - \Big(\frac{\xi}{4}\Big)^2}\,\Big] = \frac{1}{4} + \frac{\tilde\eta}{4\tilde\xi^2}\Big[1-\sqrt{1-\tilde\xi^2}\Big]\,, \end{aligned}$$

where $\tilde\xi = \xi/(4-\eta)$, $\tilde\eta = \eta/(4-\eta)$. Since $\xi \le \eta \le 4-\eta$ and $0 \le 1-\sqrt{1-\tilde\xi^2} \le \tilde\xi^2$ it follows that $1/4 \le f(w,z) \le 1/4 + \tilde\eta/4 \le 1/2$ as needed. □

The control of Lemma 2.1 on $g(x,y)$, hence on $c_n(x,y)$, shall give rise to the perturbed centered Gaussian processes $Y^{(\alpha)}$ of the next lemma.

Lemma 2.2 *Let $\alpha \in [0,1]$ and define the covariance*

$$R^{(\alpha)}(\tau) = \mathrm{sech}\Big(\tau/2\Big)\Big\{1-\alpha(1-e^{-|\tau|})^2\Big\}\,.$$

Then there exist independent, stationary centered Gaussian processes Y_t, Z_t, with covariances $R_y(\tau) = R^{(0)}(\tau)$ and

$$R_z(\tau) = R^{(1)}(\tau) = \mathrm{sech}\Big(\tau/2\Big)\Big(2e^{-|\tau|} - e^{-2|\tau|}\Big)$$

respectively, such that the process $Y_t^{(\alpha)} := \sqrt{1-\alpha}\,Y_t + \sqrt{\alpha}Z_t$ has covariance $R^{(\alpha)}(\tau)$.

Proof: Since $R^{(\alpha)}(\tau) = (1-\alpha)R_y(\tau) + \alpha R_z(\tau)$, all one needs is to check that both $R_y(\tau)$ and $R_z(\tau)$ are covariance functions, i.e. to check that their Fourier transforms are nonnegative. To this end, note that

$$S_y(\omega) := \mathcal{F}(R_y(\tau)) = \int_{-\infty}^{\infty} e^{i\omega\tau} R_y(\tau)d\tau = 2\int_0^\infty \cos(\omega\tau)\mathrm{sech}(\tau/2)d\tau = 2\pi\mathrm{sech}(\omega\pi) \ge 0\,,$$

c.f. [GR, p. 503, formula 3.981.3]. Furthermore,

$$S_z(\omega) := \mathcal{F}(R_z(\tau)) = \int_{-\infty}^{\infty} e^{i\omega\tau} R_z(\tau)d\tau = S_y(\omega) * F(\omega)\,,$$

where $*$ stands throughout for the convolution operation and

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} e^{i\omega\tau}\Big(2e^{-|\tau|} - e^{-2|\tau|}\Big)d\tau = 2\int_0^\infty \cos(\omega\tau)\Big(2e^{-\tau} - e^{-2\tau}\Big)d\tau \\ &= \frac{4}{1+\omega^2} - \frac{4}{4+\omega^2} = \frac{12}{(4+\omega^2)(1+\omega^2)} \ge 0\,. \end{aligned}$$

Hence, $S_z(\omega) \ge 0$. □

The effect of nonrandom functions $\gamma_n(x)$ as well as that of considering the processes $Y^{(\alpha)}$ for some $\alpha_n \downarrow 0$ are dealt with by the continuity properties of Y_t and Z_t outlined in the next lemma.

Lemma 2.3 *Let Y_t, Z_t be as in Lemma 2.2. Then, for any positive $\varepsilon_T \to 0$,*

$$\lim_{T\to\infty} P\Big(\sup_{0\le t\le T} Z_t \le \sqrt{\varepsilon_T^{-1}\log T}\Big) = 1\,, \tag{2.4}$$

whereas

$$\begin{aligned}\limsup_{T\to\infty}\frac{1}{T}\log P\Big(\sup_{0\le t\le T} Y_t \le \varepsilon_T\Big) &= \liminf_{T\to\infty}\frac{1}{T}\log P\Big(\sup_{0\le t\le T} Y_t \le -\varepsilon_T\Big) \\ &= \lim_{T\to\infty}\frac{1}{T}\log P\Big(\sup_{0\le t\le T} Y_t \le 0\Big) = -\frac{b}{4}. \qquad (2.5)\end{aligned}$$

Moreover, for any positive $\gamma_T \to 0$ *and* $\alpha_T \log T \to 0$,

$$\liminf_{T\to\infty}\frac{1}{T}\log P\Big(\inf_{0\le t\le T} Y_t^{(\alpha_T)} \ge \gamma_T\Big) \ge -\frac{b}{4}. \qquad (2.6)$$

Proof: The existence of the limit in the right hand side of (2.5) (and hence in (1.3)) is ensured by sub-additivity: since $R_y(\cdot) > 0$, Slepian's lemma (c.f. [Ad, Page 49]), and the stationarity of $Y.$ imply

$$\begin{aligned}P\Big(\sup_{0\le t\le T+S} Y_t \le 0\Big) &\ge P\Big(\sup_{0\le t\le T} Y_t \le 0\Big)P\Big(\sup_{T\le t\le T+S} Y_t \le 0\Big) \\ &= P\Big(\sup_{0\le t\le T} Y_t \le 0\Big)P\Big(\sup_{0\le t\le S} Y_t \le 0\Big).\end{aligned}$$

Fix $\varepsilon_T \to 0$. From Lemma 2.2, we have that

$$S_z(\omega) = \mathcal{F}(R_z(\tau)) = 2\pi\mathrm{sech}(\omega\pi) * \frac{12}{(4+\omega^2)(1+\omega^2)},$$

which implies that $\sup_\omega\{S_z(\omega)\omega^4\} < \infty$. Hence $\int_{-\infty}^{\infty}\omega^2 S_z(\omega)d\omega < \infty$. It follows that

$$-\frac{\partial^2}{\partial\tau^2}R_z(\tau)\Big|_{\tau=0} = E(\dot{Z}_t^2) < \infty.$$

Since $|Z_t| \le |Z_0| + \int_0^1 |\dot{Z}_t|dt$, it follows by stationarity of the centered Gaussian process $\dot{Z}_t$ that

$$m_1 := E(\sup_{0\le t\le 1}|Z_t|) \le \sqrt{E(Z_0^2)} + \sqrt{E(\dot{Z}_t^2)} < \infty.$$

By the stationarity of Z_t and Borell's inequality (c.f. [Ad, Page 43]), for all $\lambda \ge m_1$,

$$P\Big(\sup_{0\le t\le T} Z_t \ge \lambda\Big) \le TP\Big(\sup_{0\le t\le 1}|Z_t| \ge \lambda\Big) \le 2T\exp\Big(-\frac{(\lambda-m_1)^2}{2R_z(0)}\Big).$$

Setting $\lambda = \sqrt{\varepsilon_T^{-1}\log T}$ we obtain that as $T\to\infty$,

$$P\Big(\sup_{0\le t\le T} Z_t \ge \sqrt{\varepsilon_T^{-1}\log T}\Big) \to 0,$$

which yields (2.4).

To see (2.5), let the Gaussian law of $Y.$ on $C(\mathbb{R};\mathbb{R})$ be denoted by P_y. Let $\mathcal{R}_y$ denote the covariance operator associated with P_y, that is, $\mathcal{R}_y g(t) = \int_{-\infty}^{\infty}\mathcal{R}_y(t-s)g(s)ds$, with $K_y = \mathcal{R}_y^{-1}$ denoting its inverse (defined on the range of $\mathcal{R}_y$). We also let $\langle\cdot,\cdot\rangle$ denote the inner product of $L^2(\mathbb{R})$. Fixing $T<\infty$ note that the deterministic function $f_T(t) := \varepsilon_T\exp(\frac{1}{4} - \left(\frac{t}{T}\right)^2)$ is in the Reproducing Kernel Hilbert Space (RKHS) associated with the process $Y.$. Indeed, the Fourier transform of f_T is $\hat{f}(\omega) = c_1 T \varepsilon_T e^{-\omega^2 T^2}$ (for some $c_1 < \infty$), so it follows by Parseval's theorem that for some $c_2 < \infty$ and all T,

$$\langle f_T, K_y f_T \rangle = \int_{-\infty}^{\infty} \frac{|\hat{f}(\omega)|^2}{S_y(\omega)} d\omega = 2\pi \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 \mathrm{sech}(\omega\pi) d\omega \leq c_2 T \varepsilon_T^2 \,, \tag{2.7}$$

In particular, $\langle f_T, K_y f_T \rangle$ is finite and the Radon-Nikodým derivative

$$\Lambda_T(Y) = \exp\Big(\langle f_T, K_y Y \rangle - \frac{1}{2} \langle f_T, K_y f_T \rangle\Big) \,,$$

is well defined and finite for P_y-almost-every $Y_\cdot$. Since $f_T(t) \geq \varepsilon_T$ for all $-\frac{T}{2} \leq t \leq \frac{T}{2}$, it follows that

$$\begin{aligned} & P\Big(\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} Y_t \leq \varepsilon_T\Big) = P\Big(\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} \{Y_t - \varepsilon_T\} \leq 0\Big) \leq P\Big(\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} \{Y_t - f_T(t)\} \leq 0\Big) \\ = \;& E\Big(\Lambda_T(Y)\, \mathbf{1}_{\{\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} Y_t \leq 0\}}\Big) \leq E(\Lambda_T(Y)^q)^{\frac{1}{q}} \Big[P\Big(\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} Y_t \leq 0\Big)\Big]^{\frac{1}{p}} \end{aligned} \tag{2.8}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Note that

$$\Big(E(\Lambda_T(Y)^q)\Big)^{\frac{1}{q}} = \exp\Big(\frac{q-1}{2} \langle f_T, K_y f_T \rangle\Big) \,.$$

Hence, choosing $q_T = (1/\varepsilon_T) \to \infty$ it follows from (2.7) that

$$\frac{1}{T} \log\Big(E(\Lambda_T(Y)^q)\Big)^{\frac{1}{q}} \underset{T \to \infty}{\longrightarrow} 0 \,.$$

Substituting in (2.8) and using the stationarity of $Y_\cdot$ and existence of the limit in (1.3), one has that

$$\limsup_{T \to \infty} \frac{1}{T} \log P\Big(\sup_{0 \leq t \leq T} Y_t \leq \varepsilon_T\Big) \leq \lim_{T \to \infty} \frac{1}{T} \log P\Big(\sup_{0 \leq t \leq T} Y_t \leq 0\Big) \,. \tag{2.9}$$

The equality in (2.9) is then obvious. The other equality in (2.5) follows by a similar proof, starting with

$$P\Big(\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} Y_t \leq 0\Big) \leq P\Big(\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} \{Y_t - f_T(t)\} \leq -\varepsilon_T\Big) \leq E(\Lambda_T(Y)^q)^{\frac{1}{q}} P\Big(\sup_{-\frac{T}{2} \leq t \leq \frac{T}{2}} Y_t \leq -\varepsilon_T\Big)^{\frac{1}{p}} .$$

Turning to prove (2.6), set $\varepsilon_T = 3\max(\gamma_T, (\alpha_T \log T)^{1/3}) \to 0$, and note that

$$\sqrt{1 - \alpha_T} \varepsilon_T - \sqrt{\alpha_T} \sqrt{\varepsilon_T^{-1} \log T} \geq \gamma_T \,,$$

once T is large enough that $\alpha_T \leq 1/3$. Then, by the independence of Y_t and Z_t,

$$P\Big(\inf_{0 \leq t \leq T} Y_t^{(\alpha_T)} \geq \gamma_T\Big) \geq P\Big(\inf_{0 \leq t \leq T} Y_t \geq \varepsilon_T\Big) P\Big(\inf_{0 \leq t \leq T} Z_t \geq -\sqrt{\varepsilon_T^{-1} \log T}\Big) \,.$$

With the laws of the processes Y_t and Z_t invariant to a change of sign, the inequality (2.6) is thus a direct consequence of (2.4) and (2.5). □

The control of $\hat{f}_n^b(x)$ for $x \in [1 - n^{-1}, 1]$ is achieved in the next lemma by means of the sample path smoothness of $f_n^b(\cdot)$.

Lemma 2.4 *For any finite γ, the set of limit points of $C_n = P(\hat{f}_n^b(x) > \gamma, \forall x \in [1-n^{-1}, 1])$ is bounded below by some $C_\infty = C_\infty(\gamma) > 0$.*

Proof: Without loss of generality we assume that $\gamma \geq 0$. Since $x \mapsto E(f_n^b(x)^2)$ is increasing on $[0, \infty)$, with $E(f_n^b(1)^2) = n$, it follows that for any $\lambda > 0$

$$\begin{aligned} C_n &\geq P\Big(f_n^b(x) > \gamma\sqrt{n}, \quad \forall x \in [1-n^{-1}, 1]\Big) \\ &\geq P\Big(f_n^b(1) > (\lambda+\gamma)\sqrt{n}\Big) - P\Big(\sup_{(1-n^{-1})\leq\xi\leq 1} f_n^b{}'(\xi) \geq \lambda n^{3/2}\Big) \end{aligned} \tag{2.10}$$

We wish to apply Borell's inequality to bound the second term in (2.10). To this end, note that

$$f_n^b{}'(\xi) = \sum_{i=0}^n i b_i \xi^{i-1} = \sum_{i=0}^n \Big(\sum_{j=0}^i b_j\Big)\Big[i\xi^{i-1} - (i+1)\xi^i\Big] + (n+1)\xi^n \sum_{j=0}^n b_j \,.$$

By Kolmogorov's maximal inequality,

$$E\Big[\sup_{i\leq n}\Big|\sum_{j=0}^i b_j\Big|\Big] \leq c_1 n^{1/2}\,.$$

Hence, for some $c_2 > 0$,

$$E\Big|\sup_{1-n^{-1}\leq\xi\leq 1} f_n^b{}'(\xi)\Big| \leq c_2 n^{3/2}\,. \tag{2.11}$$

Furthermore, we have that

$$\sup_{1-n^{-1}\leq\xi\leq 1} E\Big[\frac{1}{n^{3/2}} f_n^b{}'(\xi)\Big]^2 = \frac{1}{n^3}\sum_{i=1}^{n-1} i^2 \underset{n\to\infty}{\longrightarrow} \frac{1}{3}\,,$$

implying, by Borell's inequality and (2.11), that for some finite c_3, all n and any $\lambda \geq c_2$,

$$P\Big(\sup_{1-n^{-1}\leq\xi\leq 1} f_n^b{}'(\xi) \geq \lambda n^{3/2}\Big) \leq c_3 e^{-3(\lambda-c_2)^2/2}\,. \tag{2.12}$$

Since $n^{-1/2} f_n^b(1)$ is a standard Normal random variable, it follows that for some positive $c_4 = c_4(\gamma)$, $\lambda = \lambda(\gamma)$ large enough and all n,

$$P\Big(f_n^b(1) > (\lambda+\gamma)\sqrt{n}\Big) \geq c_4 e^{-\lambda^2} \geq 2c_3 e^{-3(\lambda-c_2)^2/2}\,. \tag{2.13}$$

Substituting (2.12) and (2.13) in (2.10), one concludes that $\liminf_{n\to\infty} C_n \geq C_\infty > 0$ as claimed. $\square$

The next lemma provides the bounds on the value of b stated in the introduction.

Lemma 2.5 *The limit in (1.3) exists, and the constant b there satisfies the bounds $0.4 \leq b \leq 2$.*

Proof: The existence of the limit in (1.3) was proved in the course of proving Lemma 2.3. Recall that $R_y(t) \geq e^{-|t|/2}$, the covariance of the stationary Ornstein-Uhlenbeck process X. As can be checked by computing the covariance, a representation of the process $\{X_t\}$ can be obtained as

$$X_t = e^{-t/2} V_{e^t} = e^{-t/2}(V_{e^t} - V_1 + V_1) = e^{-t/2}(W_{e^t-1} + X_0)\,, \tag{2.14}$$

for some standard Brownian motions $V_\cdot$, $W_\cdot$ and a standard normal random variable X_0 that is independent of $W_\cdot$. Hence, for $\eta = (e^T - 1)^{-1/2}$,

$$\begin{aligned} P(\sup_{0\le t\le T} X_t \le 0) &= E[1_{X_0\le 0}\, P(\sup_{0\le t\le e^T-1} \{W_t\} \le -X_0 | X_0)] \\ &= E[1_{X_0\le 0}\,(1 - 2P(W_{e^T-1} \ge -X_0|X_0))] \\ &= \frac{1}{\pi}\int_{-\infty}^{0}\int_{0}^{-\eta x} e^{-(x^2+y^2)/2} dy dx = \frac{1}{\pi}\arctan(\eta) = \frac{1}{\pi} e^{-T/2}(1+o(1)). \end{aligned}$$

Consequently, Slepian's lemma implies the bound $b \le 2$.

The proof of the complementary bound is based on the following observation. Suppose that $\mathbf{X} \in \mathbb{R}^n$ and $\mathbf{Y} \in \mathbb{R}^n$ are zero-mean, normally distributed random vectors with covariance matrices Σ_x and Σ_y respectively. If $\Sigma_x - \Sigma_y$ is a positive semidefinite matrix, then the Radon-Nikodým derivative of the law of $\mathbf{Y}$ with respect to that of $\mathbf{X}$ is at most $\big(\det \Sigma_x / \det \Sigma_y\big)^{1/2}$, hence

$$P(\mathbf{Y} \in C) \le \Big(\frac{\det \Sigma_x}{\det \Sigma_y}\Big)^{1/2} P(\mathbf{X} \in C), \tag{2.15}$$

for all $C \subset \mathbb{R}^n$ (c.f. [Sh, Lemma 3.1]). Indeed, to prove that $b \ge 0.4$, it suffices to show that

$$P(\max_{1\le i\le n} Y_{5i} \le 0) \le \exp(-0.5\,n) \tag{2.16}$$

for all $n \ge 2$. Let

$$\rho = 2\,\mathrm{sech}(5/2) = \frac{2e^{-2.5}}{1+e^{-5}}, \quad \lambda_0 = 4\sum_{i=2}^{\infty} \mathrm{sech}(5i/2) = 4\sum_{i=2}^{\infty} \frac{e^{-2.5i}}{1+e^{-5i}}$$

and $(X_1, \cdots, X_n)$ be independent normal random variables each having zero mean and variance $\lambda := 1 + 2\rho + \lambda_0$. Denote the covariance matrices of $(X_i, 1 \le i \le n)$ and $(Y_{5i}, 1 \le i \le n)$ by Σ_x and Σ_y, respectively. It is easy to see that $\Sigma_x - \Sigma_y$ is a dominant principal diagonal matrix and as such it is positive semidefinite. Thus, by (2.15)

$$P(\max_{1\le i\le n} Y_{5i} \le 0) \le \Big(\frac{\det \Sigma_x}{\det \Sigma_y}\Big)^{1/2} P(\max_{1\le i\le n} X_i \le 0) = \frac{\lambda^{n/2} 2^{-n}}{(\det \Sigma_y)^{1/2}}.$$

To estimate $\det \Sigma_y$, let $\Sigma_n = (r_{ij}, 1 \le i,j \le n)$ be a tri-diagonal matrix with $r_{ii} = 1 - \lambda_0$, $r_{i,i+1} = r_{i-1,i} = \rho$ and $r_{ij} = 0$ for other i,j. Then, $\Sigma_y - \Sigma_n$ is a positive semidefinite matrix and hence

$$\det \Sigma_y \ge \det \Sigma_n := D_n.$$

Since $D_n = (1-\lambda_0) D_{n-1} - \rho^2 D_{n-2}$, direct calculation shows that

$$D_n \ge \Big(\frac{1}{2}(1 - \lambda_0 + \sqrt{(1-\lambda_0)^2 - 4\rho^2})\Big)^n.$$

Putting the above inequalities together yields

$$\begin{aligned} P(\max_{1\le i\le n} Y_{5i} \le 0) &\le \left(\frac{\lambda}{2(1-\lambda_0 + \sqrt{(1-\lambda_0)^2 - 4\rho^2})}\right)^{n/2} \\ &= \exp\Big\{0.5n \ln\Big(\frac{\lambda}{2(1-\lambda_0 + \sqrt{(1-\lambda_0)^2 - 4\rho^2})}\Big)\Big\} \\ &\le \exp(-0.5n) \end{aligned}$$

(here, $\lambda = 1.3555\cdots$, $\rho = 0.163071\cdots$ and $\lambda_0 = 0.029361\cdots$). □

3 Lower bound for Theorem 1.4

Hereafter let $\theta_1(x) = x$, $\theta_2(x) = x^{-1}$, $\theta_3(x) = -x^{-1}$ and $\theta_4(x) = -x$ be the symmetry transformations preserved by the Gaussian processes $\hat{f}_n^b(x)$ and let $\overline{\gamma}_n(x) = \max_{j=1}^4 \gamma_n(\theta_j(x))$ (with the exception of $x = 0$ for which $\overline{\gamma}_n(0) = \gamma_n(0)$). We begin by noting that, with $I_1 = [0,1]$, $I_2 = [1,\infty)$, $I_3 = (-\infty,-1]$ and $I_4 = [-1,0]$,

$$\begin{aligned} P(\hat{f}_n^b(x) > \gamma_n(x),\, \forall x \in \mathbb{R}) &= P(\hat{f}_n^b(x) > \gamma_n(x),\, \forall x \in I_1 \cup I_2 \cup I_3 \cup I_4) \\ &\geq \prod_{i=1}^4 P(\hat{f}_n^b(x) > \gamma_n(x),\, \forall x \in I_i) \\ &\geq \Big[P(\hat{f}_n^b(x) > \overline{\gamma}_n(x),\, \forall x \in [0,1])\Big]^4 \end{aligned} \tag{3.1}$$

where the first inequality follows by Slepian's lemma due to the positivity of the covariance $c_n(x,y)$ of $\hat{f}_n^b$, while the second holds because $c_n(x,y) = c_n(-x,-y) = c_n(\frac{1}{x},\frac{1}{y})$. Set $T = \log n$. The assumptions of Theorem 1.4 imply the existence of the integers $\log\log T \ll \tau_T \ll T$ such that $\delta_n := \sup\{\overline{\gamma}_n(x) : x \in [1-\xi_n, 1]\} \to 0$ for $\xi_n = \exp(-\tau_T)$. Recall also our assumption that $\sup\{\gamma_n(x) : x \in \mathbb{R}, n\} \leq M < \infty$. Applying Slepian's lemma once more yields that,

$$\begin{aligned} & P(\hat{f}_n^b(x) > \overline{\gamma}_n(x),\, \forall x \in [0,1]) \\ \geq\ & P\Big(\inf_{0\leq x\leq 1-\xi_n} \hat{f}_n^b(x) > M\Big) P\Big(\inf_{1-\xi_n\leq x\leq 1-n^{-1}} \hat{f}_n^b(x) > \delta_n\Big) P\Big(\inf_{1-n^{-1}\leq x\leq 1} \hat{f}_n^b(x) > M\Big) \\ :=\ & A_n B_n C_n\,. \end{aligned} \tag{3.2}$$

Starting with A_n, note that for $1 > x \geq y \geq 0$ one has

$$1 \leq g(x,y) = \frac{1-xy}{\sqrt{(1-x^2)(1-y^2)}} \leq \sqrt{\frac{1-y}{1-x}}$$

and hence, by (2.2), taking $x = 1-e^{-t}$ and $y = 1-e^{-s}$ we see that for $x, y \in [0,1)$,

$$c_n(x,y) \geq \sqrt{\frac{1-x\vee y}{1-x\wedge y}} = e^{-|t-s|/2}\,. \tag{3.3}$$

Recall that $\exp(-|t-s|/2)$ is the covariance of the stationary Ornstein-Uhlenbeck process (see (2.14)). In view of (3.3), we have by Slepian's lemma that

$$A_n = P\Big(\inf_{0\leq x\leq 1-\xi_n} \hat{f}_n^b(x) > M\Big) \geq P\Big(\inf_{0\leq t\leq \tau_T} X_t > M\Big)\,.$$

Since X_t is a centered stationary Gaussian process of positive covariance, yet another application of Slepian's lemma yields that

$$\liminf_{n\to\infty} \frac{\log A_n}{\log n} \geq \liminf_{T\to\infty} \frac{\tau_T}{T} \log P(\inf_{0\leq t\leq 1} X_t > M) = 0\,. \tag{3.4}$$

(since the random variable $\inf_{0\leq t\leq 1} X_t$ is unbounded).

We next turn to the dominant term B_n. Setting $z = 1-x, w = 1-y$, for all $x, y \in [1-\xi_n, 1)$, n large, it follows from (2.2), (2.3) and Lemma 2.1 that

$$c_n(x,y) \geq \frac{1}{g(x,y)} \geq \frac{2\sqrt{zw}}{z+w}\Big[1 - \frac{(z-w)^2}{\max(z,w)}\Big]$$

Making yet another change of variables $z = e^{-t}, w = e^{-s}$, we thus get that for $\alpha = e^{-\tau_T}$, in the notations of Lemma 2.2,

$$c_n(x,y) \geq \frac{2e^{-\frac{|s-t|}{2}}}{1+e^{-|s-t|}}\Big[1-\alpha(1-e^{-|s-t|})^2\Big] = R^{(\alpha)}(s-t)\,.$$

With $R^{(\alpha)}(0) = c_n(x,x) = 1$, it follows by Slepian's lemma that

$$B_n \geq P\Big(\inf_{0\leq t\leq T} Y_t^{(\alpha)} > \delta_n\Big)$$

Since $\delta_n \to 0$ and $\alpha_T \log T \to 0$ by our choice of τ_T, it follows by (2.6) of Lemma 2.3 that

$$\liminf_{n\to\infty} \frac{\log B_n}{\log n} \geq -\frac{b}{4}\,. \tag{3.5}$$

Finally, we recall that the sequence C_n is bounded away from zero by Lemma 2.4. Combining (3.1), (3.2), (3.4) and (3.5), we thus arrive at the stated lower bound

$$\liminf_{n\to\infty} \frac{\log P^b_{n,\gamma_n}}{\log n} \geq -b\,,$$

of (1.9). $\square$

4 Upper bound for Theorem 1.4

Fixing $\frac{1}{2} > \delta > 0$, define the four disjoint intervals $\mathcal{I}_1 = [1-n^{-\delta}, 1-n^{-(1-\delta)}]$ and $\mathcal{I}_j = \theta_j(\mathcal{I}_1)$, $j = 2,3,4$. Let $V = \bigcup_{j=1}^4 \mathcal{I}_j$ and $U = \bigcup_{j=1}^4 \{(x,y) : x,y \in \mathcal{I}_j\}$.

The crucial tool in the proof of the upper bound is the following lemma, whose proof is deferred to the end of this section:

Lemma 4.1 *For all n sufficiently large there exist $0 \leq \alpha_n \leq n^{-\delta/2}$ such that*

$$c_n(x,y) \leq \frac{(1-\alpha_n)}{g(x,y)} \mathbf{1}_{(x,y)\in U} + \alpha_n \qquad \forall x,y \in V\,. \tag{4.1}$$

Equipped with Lemma 4.1, we show how to complete the proof of the upper bound. Let $\{N, b_i^{(j)}, j = 1,2,3,4, i = 0,\ldots\}$ be independent, identically distributed standard Normal random variables. For $x \in \mathcal{I}_1$ consider the infinite random polynomials $\hat{f}_\infty^{(j)}(x) = \sqrt{1-x^2}\sum_{i=0}^\infty b_i^{(j)} x^i$ which are for $j = 1,2,3,4$ well defined i.i.d. centered Gaussian processes of covariance function $1/g(x,y)$. Recall that $g(x,y)$ is invariant to application of each of the invertible transformations $\theta_j(\cdot)$, $j = 2,3,4$ on both x and y. Each such transformation is a one to one map of $\mathcal{I}_j$ to $\mathcal{I}_1$. Hence, the right hand side of (4.1) represents the covariance of the centered Gaussian field $\widetilde{f}_n(\cdot)$ defined on V, of the form

$$\widetilde{f}_n(x) = \sqrt{1-\alpha_n}\sum_{j=1}^4 \mathbf{1}_{x\in\mathcal{I}_j} \hat{f}_\infty(\theta_j(x)) + \sqrt{\alpha_n} N$$

Observe that the assumptions of Theorem 1.4 imply that $\eta_n := n^{-\delta/8} \vee \sup\{-\gamma_n(x) : x \in V\}$ decay to zero as $n \to \infty$. With $g(x,x) = 1$ for all $x \in V$, relying upon (4.1) and the positivity of $1/g(x,y)$

we get by two applications of Slepian's lemma that for all n sufficiently large

$$\begin{aligned}
P^b_{n,\gamma_n} &= P\Big(\hat{f}^b_n(x) < -\gamma_n(x)\,, \forall x \in \mathbb{R}\Big) \le P\Big(\sup_{x\in V} \hat{f}^b_n(x) \le \eta_n\Big) \le P\Big(\sup_{x\in V} \tilde{f}_n(x) \le \eta_n\Big) \\
&\le P\Big(N \le -n^{\delta/8}\Big) + \prod_{j=1}^{4} P\Big(\sup_{x\in\mathcal{I}_j} \hat{f}_\infty(\theta_j(x)) \le \frac{2\eta_n}{\sqrt{1-\alpha_n}}\Big) \\
&\le e^{-n^{\delta/4}/2} + P\Big(\sup_{x\in\mathcal{I}_1} \hat{f}_\infty(x) \le 3\eta_n\Big)^4 . \qquad (4.2)
\end{aligned}$$

Hence, it is enough to show that

$$\limsup_{\delta\to 0}\limsup_{n\to\infty} \frac{1}{\log n} \log P\Big(\sup_{x\in\mathcal{I}_1} \hat{f}_\infty(x) \le 3\eta_n\Big) \le -\frac{b}{4} . \qquad (4.3)$$

The change of variables $x = 1 - z = 1 - e^{-t}$, $y = 1 - w = 1 - e^{-s}$ yields, by (2.3) and Lemma 2.1, that for all sufficiently large n and all $x, y \in \mathcal{I}_1$,

$$\frac{1}{g(x,y)} \le \frac{2\sqrt{zw}}{z+w} = \operatorname{sech}\Big(\frac{t-s}{2}\Big) . \qquad (4.4)$$

For $T = \log n$, $T' = (1-2\delta)T$ and $\varepsilon_{T'} := 3\eta_n \to 0$, by yet another application of Slepian's lemma and the stationarity of the process Y_t of Lemma 2.2, it follows from (4.4) that

$$P\Big(\sup_{x\in\mathcal{I}_1} \hat{f}_\infty(x) \le 3\eta_n\Big) \le P\Big(\sup_{t\in[\delta T,(1-\delta)T]} Y_t \le 3\eta_n\Big) \le P\Big(\sup_{0\le t\le T'} Y_t \le \varepsilon_{T'}\Big) . \qquad (4.5)$$

Consequently, by (2.5),

$$\limsup_{n\to\infty} \frac{1}{\log n} \log P\Big(\sup_{x\in\mathcal{I}_1} \hat{f}_\infty(x) \le 3\eta_n\Big) \le (1-2\delta) \limsup_{T\to\infty} \frac{1}{T} \log P\Big(\sup_{0\le t\le T} Y_t \le \varepsilon_T\Big) = -(1-2\delta)\frac{b}{4}. \qquad (4.6)$$

Taking $\delta \to 0$, we see that (4.6) implies (4.3) hence the proof of the upper bound in (1.9), modulo the proof of Lemma 4.1 which we provide next.

Proof of Lemma 4.1 Considering separately $(x,y) \in U$ and $(x,y) \notin U$, it is enough by the symmetry relations $g(x,y) = g(-x,-y) = g(\frac{1}{x}, \frac{1}{y})$ to show that

$$\begin{aligned}
g(x^n, y^n) &\le (1-\alpha_n) + \alpha_n g(x,y), \qquad && x, y \in \mathcal{I}_1 \qquad (4.7a) \\
g(x^n, y^n) &\le \alpha_n g(x,y), && x \in \mathcal{I}_1, y \in \mathcal{I}_j, j \ne 1 \qquad (4.7b)
\end{aligned}$$

Turning first to (4.7a), recall that $g(x,y)$ is a symmetric function, which equals 1 on the diagonal $x = y$. We thus may and shall take without loss of generality $y > x$. Fixing $x \in \mathcal{I}_1$, the change of variables $y = y(\eta) = 1 - (1-x)(1-\eta)^2$ for $\eta \in (0,1)$ then corresponds to $\eta = 1 - \sqrt{w/z}$ where $z = 1 - x$ and $w = 1 - y$. It follows from (4.4) that for all n sufficiently large,

$$g(x,y) - 1 \ge \frac{z+w}{2\sqrt{zw}} - 1 = \frac{1}{2}\Big(1 - \sqrt{\frac{w}{z}}\Big)^2 = \frac{\eta^2}{2} .$$

Moreover, when n is large enough,

$$\sqrt{1-x^{2n}}\sqrt{1-y^{2n}} \ge \frac{2}{3} \qquad \forall x, y \in \mathcal{I}_1 \qquad (4.8)$$

So that,

$$\frac{g(x^n,y^n)-1}{g(x,y)-1} \le \frac{2}{\eta^2}\,\frac{1-x^ny^n-\sqrt{1-x^{2n}}\,\sqrt{1-y^{2n}}}{\sqrt{1-x^{2n}}\,\sqrt{1-y^{2n}}} \le \frac{3}{\eta^2}h(\eta)\,, \tag{4.9}$$

where

$$h(\eta) := 1-x^ny(\eta)^n-\sqrt{1-x^{2n}}\,\sqrt{1-y(\eta)^{2n}}\,.$$

Note that $y(0)=x$, hence $h(0)=0$. It is not hard to check that $h'(0)=0$ and

$$\begin{aligned} h''(\xi) &= ny''(\xi)\Big[\frac{\sqrt{1-x^{2n}}}{\sqrt{1-y^{2n}}}y^{2n-1}-x^ny^{n-1}\Big] \\ &\quad +ny'(\xi)^2\Big[(2n-1)y^{2n-2}\frac{\sqrt{1-x^{2n}}}{\sqrt{1-y^{2n}}}+\frac{ny^{4n-2}}{1-y^{2n}}\frac{\sqrt{1-x^{2n}}}{\sqrt{1-y^{2n}}}-(n-1)x^ny^{n-2}\Big]\,, \end{aligned}$$

evaluated at $y=y(\xi)$. Observing that $|y'(\xi)|\le 2$, $y''(\xi)\in[-2,0]$ and $x^n\le y(\xi)^n\le y(\eta)^n\le e^{-n^\delta}$ for all $\xi\in[0,\eta]$, it is easy to check that there exists a universal finite constant c_1 such that

$$\sup_{\xi\in[0,\eta]} h''(\xi)\le c_1y(\eta)^n\le c_1e^{-n^\delta}\,,$$

for all n large enough and any $x,y\in\mathcal{I}_1$. Hence, $h(\eta)\le\frac{1}{2}c_1e^{-n^\delta}\eta^2$. Substituting in (4.9), we conclude that

$$\sup_{x,y\in\mathcal{I}_1}\frac{g(x^n,y^n)-1}{g(x,y)-1}\le 2c_1e^{-n^\delta}$$

proving (4.7a).

Turning to the proof of (4.7b) we assume first that $x\in\mathcal{I}_1$ and $y\in\mathcal{I}_2\cup\mathcal{I}_3$. Then, $x,|y|^{-1}\in\mathcal{I}_1$ with $\sqrt{1-x^2}\sqrt{1-y^{-2}}\le 1$ and (4.8) holding for x and y^{-1}. Moreover, $x^n\vee|y|^{-n}\le e^{-n^\delta}$, so we have in this case that

$$\frac{g(x^n,y^n)}{g(x,y)}\le\frac{3}{2}\frac{|x^n-y^{-n}|}{|x-y^{-1}|}\le\frac{3}{2}\sum_{k=0}^{n-1}x^k|y|^{-(n-1-k)}\le 2ne^{-n^\delta}\,.$$

In the remaining case of $x\in\mathcal{I}_1$ and $y\in\mathcal{I}_4$ we have that $|y|^n\le e^{-n^\delta}$, hence $g(x^n,y^n)\le 2$ while

$$\frac{1}{g(x,y)}=\frac{\sqrt{1-x^2}\sqrt{1-y^2}}{1+|xy|}\le[1-(1-n^{-\delta})^2]\le 2n^{-\delta},$$

thus completing the proof of (4.7b). □

5 Proof of Theorem 1.3

Our proof of Theorem 1.3 combines the Komlós-Major-Tusnády strong approximation theorem with Theorem 1.4. To this end, note that for every k and $|x|\le 1$, the sequence $\{(1-x^2)x^{2j}: j=0,\dots,k-1\}\cup\{x^{2k}\}$ is a probability distribution, hence for any real valued s_j,

$$\Big|s_0+\sum_{j=1}^{k}(s_j-s_{j-1})x^{2j}\Big|=\Big|s_kx^{2k}+(1-x^2)\sum_{j=0}^{k-1}s_jx^{2j}\Big|\le\max_{0\le j\le k}|s_j|\,. \tag{5.1}$$

Recall that $E(a_i)=0$ and $E(a_i^2)=1$. Hence, applying the strong approximation theorem of [KMT] twice we can redefine $\{a_i, 0\le i\le n-1\}$ on a new probability space with a sequence of independent standard normal random variables $\{b_i,\ 0 \le i \le n-1\}$ such that for any $p \ge 2$, some $c_p < \infty$, all $t > 0$ and n,

$$P\left(\max_{0\le j\le (n-1)/2}\left|\sum_{i=0}^{j} a_{2i} - \sum_{i=0}^{j} b_{2i}\right| \ge t\right) + P\left(\max_{0\le j\le (n-2)/2}\left|\sum_{i=0}^{j} a_{2i+1} - \sum_{i=0}^{j} b_{2i+1}\right| \ge t\right) \le c_p n E|a_0|^p t^{-p}. \tag{5.2}$$

Let

$$g_k(x) := x^{k-1} f_k(x^{-1}) = \sum_{i=0}^{k-1} a_i x^{k-1-i} \ ,$$

and for $k \in \{1,\dots,n\}$, define

$$f_k^b(x) = \sum_{i=0}^{k-1} b_i x^i \ , \qquad\qquad g_k^b(x) = x^{k-1} f_k^b(x^{-1}) \ .$$

Let $\sigma_k(x) := \sqrt{E(f_k(x)^2)} = \sqrt{|1-x^{2k}|/|1-x^2|}$, when $|x| \neq 1$ with $\sigma_k(\pm 1) = \sqrt{E(f_k(\pm 1)^2)} = \sqrt{k}$. Define $\hat{f}_k(x) := f_k(x)/\sigma_k(x)$ and $\hat{f}_k^b(x) := f_k^b(x)/\sigma_k(x)$. As $\sigma_k(x) = \sqrt{E(g_k(x)^2)}$, we shall also use $\hat{g}_k(x) = g_k(x)/\sigma_k(x)$ and $\hat{g}_k^b(x) = g_k^b(x)/\sigma_k(x)$. Since

$$\left|\sum_{i=0}^{k} a_i x^i - \sum_{i=0}^{k} b_i x^i\right| \le \left|\sum_{j=0}^{\lfloor k/2\rfloor} (a_{2j} - b_{2j}) x^{2j}\right| + \left|\sum_{j=0}^{\lfloor (k-1)/2\rfloor} (a_{2j+1} - b_{2j+1}) x^{2j+1}\right| \tag{5.3}$$

we get from (5.2) by two applications of (5.1) (using once $s_j = \sum_{i=0}^{j}(a_{2i} - b_{2i})$ and once $s_j = \sum_{i=0}^{j}(a_{2i+1} - b_{2i+1})$) that for all $k \le n$,

$$P\left(\sup_{|x|\le 1}\left|f_k(x) - f_k^b(x)\right| \ge 2t\right) \le c_p n E|a_0|^p t^{-p} \tag{5.4}$$

The same construction of $\{b_i\}$ leads by a similar argument also to

$$P\left(\sup_{|x|\le 1}\left|g_k(x) - g_k^b(x)\right| \ge 4t\right) \le c_p n E|a_0|^p t^{-p} \tag{5.5}$$

Indeed, bounding $g_k - g_k^b$ amounts to changing (a_i, b_i) to (a_{k-1-i}, b_{k-1-i}), resulting with using once $s_j = \sum_{i=0}^{j}(a_{k-1-2i} - b_{k-1-2i})$ and once $s_j = \sum_{i=0}^{j}(a_{k-1-2i-1} - b_{k-1-2i-1})$. One controls all these as before, but for doubling the total approximation error.

In order to apply effectively the strong approximation results, we need that contributions to the value of $f_n(x)$ come from many variables. This obviously is easier for $||x|-1|$ small. In order to avoid appearance of zeros in other locations, we decompose f_n to the dominant "bulk term", which will not be too negative *everywhere* and will be rather far from 0 for $||x|-1|$ small, and to "boundary terms", which involve a small number of coefficients and thus can be made to have prescribed positive values with a not too small probability.

In order to define precisely the different regions considered for values of x and the splitting into bulk and boundary terms, we introduce, for n large enough odd integers, a few n-dependent parameters as follows:

p_n :	$p_n \uparrow \infty, c_{p_n} E\|a_0\|^{p_n} \le n$	c_p is the KMT constant in (5.2).
$\epsilon = \epsilon_n$:	$\epsilon_n \downarrow 0, \epsilon_n \ge \max\{20/p_n, (\log n)^{-1/2}\}$, $2n^{3\epsilon_n} = 2^j$ for some j integer	ϵ_n is taken as the smallest possible value satisfying constraints.
$m = m_n$:	$m_n \to \infty, m_n = 2n^{3\epsilon_n}$	m_n is an integer power of 2.
$\bar{\gamma}_n(x)$:	$\bar{\gamma}_n \to 0, \bar{\gamma}_n(x) = \max\{0, \gamma_n(x), \gamma_n(x^{-1})\}$	$\gamma_n(x)$ as in statement of theorem.
ρ_n :	$\rho_n \to 0, \rho_n = \sup_{\|x\| \le 1-m^{-1}} \{\sigma_n(x)\bar{\gamma}_n(x)\}$	$\rho_n \le cn^{-\delta/2}$, some finite $c > 0$.
$r = r_n$:	$cn^{-\delta/2}$	for $n \ge 3m$; c is as in bound on ρ_n.
$\xi_n(x)$:	$\xi_n(x) = 6x^m \sigma_{n-2m}(x)\bar{\gamma}_n(x)$	

(5.6)

In order to state the decomposition alluded to above, first partition the interval $[-1,1]$ to $\mathcal{I} = \{x : |x| \ge 1 - 0.5n^{-\epsilon}\}$ and $\mathcal{I}^c = [-1,1] \setminus \mathcal{I}$. We note that $2r + \xi_n(x) \ge \sigma_n(x)\bar{\gamma}_n(x)$ for all $x \in \mathcal{I}$. Next, let $f_n = f_n^L + f_n^M + f_n^H$ where

$$f_n^L(x) = \sum_{i=0}^{m-1} a_i x^i, \qquad f_n^M(x) = \sum_{i=m}^{n-1-m} a_i x^i, \qquad f_n^H(x) = \sum_{i=n-m}^{n-1} a_i x^i \,. \tag{5.7}$$

Similarly, we let $g_n = g_n^L + g_n^M + g_n^H$ with $g_n^L(x) = x^{n-1} f_n^L(x^{-1})$, etc. With these definitions, we have the inclusions

$$\begin{aligned}
\{\hat{f}_n(x) > \gamma_n(x), \quad \forall\, x \in \mathbb{R}\} \;&\supset\; \{\hat{f}_n(x) > \bar{\gamma}_n(x),\ \hat{g}_n(x) > \bar{\gamma}_n(x), \quad \forall\, x \in [-1,1]\} \\
&\supset\; \{f_n^M(x) > \xi_n(x),\ g_n^M(x) > \xi_n(x), \quad \forall\, x \in \mathcal{I}\} \\
&\quad \cap \{f_n^M(x) > -r,\ g_n^M(x) > -r, \quad \forall\, x \in \mathcal{I}^c\} \\
&\quad \cap \{f_n^L(x) > 3r,\ g_n^L(x) \ge -r, \quad \forall x \in [-1,1]\} \\
&\quad \cap \{f_n^H(x) \ge -r,\ g_n^H(x) > 3r, \quad \forall x \in [-1,1]\}
\end{aligned}$$

(f_n^M, g_n^M are the "bulk terms" whereas $f_n^L, g_n^L, f_n^H, g_n^H$ are the "boundary terms"). Since the polynomial pairs (f_n^L, g_n^L), (f_n^M, g_n^M) and (f_n^H, g_n^H) are mutually independent, it follows that

$$\begin{aligned}
P_{n,\gamma_n} &= P\Big(\hat{f}_n(x) > \gamma_n(x), \quad \forall\, x \in \mathbb{R}\Big) \\
&\ge P\Big(\{f_n^M(x) > \xi_n(x),\ g_n^M(x) > \xi_n(x),\ \forall\, x \in \mathcal{I}\} \cap \{f_n^M(x) > -r,\ g_n^M(x) > -r,\ \forall\, x \in \mathcal{I}^c\}\Big) \\
&\quad P\Big(f_n^L(x) > 3r,\ \ g_n^L(x) \ge -r, \quad \forall\, x \in [-1,1]\Big) \\
&\quad P\Big(f_n^H(x) \ge -r,\ \ g_n^H(x) > 3r, \quad \forall\, x \in [-1,1]\Big)
\end{aligned} \tag{5.8}$$

Note that g_n^M and f_n^M are identically distributed, as are the polynomial pairs $x^{-m}(f_n^M, g_n^M)$ and (f_{n-2m}, g_{n-2m}). Thus, we have that

$$\begin{aligned}
&P\Big(\{f_n^M(x) > \xi_n(x),\ \ g_n^M(x) > \xi_n(x), \quad \forall\, x \in \mathcal{I}\} \cap \{f_n^M(x) > -r,\ \ g_n^M(x) > -r, \quad \forall\, x \in \mathcal{I}^c\}\Big) \\
&\quad \ge\ P\Big(\hat{f}_{n-2m}(x) > 6\bar{\gamma}_n(x),\ \hat{g}_{n-2m}(x) > 6\bar{\gamma}_n(x),\ \ \forall\, x \in \mathcal{I}\Big) - 2P\Big(f_n^M(x) \le -r, \text{ for some } x \in \mathcal{I}^c\Big) \\
&\quad := \ Q_1 - 2Q_2
\end{aligned}$$

Since the polynomial pairs (f_n^L, g_n^L), (g_n^H, f_n^H) and $(f_m, x^{n-m} g_m)$ have identical laws, it now follows that

$$P_{n,\gamma_n} \ge (Q_1 - 2Q_2)(Q_3 - Q_4)^2 \,,$$

where

$$Q_3 := P\Big(f_m(x) > 3r, \ \ \forall\, x \in [-1,1], \quad x^{n-m} g_m(x) \geq -r, \ \ \forall\, |x| \in [1-m^{-1},1]\Big)\,,$$

and

$$Q_4 := P\Big(x^{n-m} g_m(x) \leq -r, \text{ for some } |x| \leq 1-m^{-1}\Big).$$

To deal with the dominant term Q_1, we consider (5.4) and (5.5) for $p = p_n$ as above, $k = n-2m$, and $t = n^{\epsilon/4}$. Noting that $\eta_n = \sup\{6\bar{\gamma}_n(x) + 4t/\sigma_{n-2m}(x) : x \in \mathcal{I}\}$ approaches zero as $n \to \infty$, we get that for all n large enough,

$$\begin{aligned} Q_1 &\geq P\Big(\hat{f}^b_{n-2m}(x) > \eta_n, \quad \hat{g}^b_{n-2m}(x) > \eta_n, \quad \forall\, x \in \mathcal{I}\Big) - 2n^{-3} \\ &\geq P\Big(f^b_{n-2m}(x) > \eta_n \sigma_{n-2m}(x), \qquad \forall\, ||x|-1| \leq n^{-\epsilon_n}\Big) - 2n^{-3} \\ &\geq (n-2m)^{-b+o(1)}, \end{aligned} \tag{5.9}$$

where the last inequality follows by applying Theorem 1.4 for threshold $\eta_n \to 0$ for $||x|-1| \leq n^{-\epsilon_n}$ and zero otherwise.

Turning to estimate Q_2, recall that f^M_n has the same distribution as $x^m f_{n-2m}$ and $m = 2n^{3\epsilon}$. Recall also that $\epsilon \geq (\log n)^{-1/2}$, implying that $n^c \exp(-n^\epsilon) \to 0$ for any fixed $c < \infty$. Hence, for all n large enough,

$$\begin{aligned} Q_2 &\leq P\Big(\sup_{x\in\mathcal{I}^c} |x|^m |f_{n-2m}(x)| \geq r\Big) \\ &\leq P\Big(\sup_{x\in\mathcal{I}^c} |f_{n-2m}(x)| \geq cn^{-\delta/2}\Big(1-\frac{1}{2n^\epsilon}\Big)^{2n^{3\epsilon}}\Big) \leq P\Big(\sup_{x\in\mathcal{I}^c} |f_{n-2m}(x)| \geq 2\exp(n^\epsilon)\Big). \end{aligned}$$

Observe that for any $x, y \in [-1,1]$,

$$E\left(|f_{n-2m}(x) - f_{n-2m}(y)|^2\right) \leq \sum_{i=1}^n (x^i - y^i)^2 \leq (x-y)^2 n^3\ .$$

Recall the following well known lemma (see [Sto] for a proof).

Lemma 5.1 *Let $\{T_x, x \in [a,b]\}$ be an a.s. continuous stochastic process with $T_a = 0$. Assume that*

$$\forall\ \ x,y \in [a,b], \quad E|T_x - T_y|^2 \leq K(x-y)^2.$$

Then, we have

$$E\Big(\sup_{x\in[a,b]} T_x^2\Big) \leq 4K(b-a)^2.$$

Applying Lemma 5.1 for $T_x = f_{n-2m}(x) - f_{n-2m}(0)$, first when $x \in [0,1]$, then when $x \in [-1,0]$, we get by Markov's inequality that for all n large enough, and any $c_1 < \infty$ (for our use, $c_1 = 3$ will do),

$$Q_2 \leq P\Big(|a_0| \geq \exp(n^\epsilon)\Big) + P\Big(\sup_{|x|\leq 1} T_x^2 \geq \exp(2n^\epsilon)\Big) \leq \exp(-2n^\epsilon)(1 + E(\sup_{|x|\leq 1} T_x^2)) = o(n^{-c_1})\,. \tag{5.10}$$

Recall that $m = 2n^{\epsilon_n}$ and $\epsilon_n \to 0$, so with g_m and f_m of identical law, it follows that for all n large enough,

$$Q_4 \leq P\Big(\sup_{|x|\leq 1-m^{-1}} |x|^{n-m} |f_m(x)| \geq r\Big) \leq P\Big(\sup_{|x|\leq 1-m^{-1}} |f_m(x)| \geq 2\exp(\sqrt{n})\Big).$$

Similarly to the derivation of (5.10), by twice applying Lemma 5.1 for $T_x = f_m(x) - f_m(0)$, then using Markov's inequality, we get that $Q_4 \le \exp(-n^{1/3})$ for all n large enough. The lower bound $P_{n,\gamma_n} \ge n^{-b+o(1)}$ in Theorem 1.3 is thus a direct consequence of the bounds (5.9), (5.10) and Lemma 5.2 below which provides the estimate $Q_3 \ge n^{-c_2\epsilon}$ with $\epsilon = \epsilon_n \to 0$ and $c_2 < \infty$ fixed.

Turning to the upper bound $P_{n,\gamma_n} \le n^{-b+o(1)}$ in Theorem 1.3, let $\eta_n := \inf\{\gamma_n(x) : ||x|-1| \le n^{-\epsilon}\}$. Recall that $\eta_n \to 0$ by our assumptions. Then, similarly to the derivation of (5.9), now with $m = 0$, we see that for all n large enough

$$\begin{aligned} P_{n,\gamma_n} &= P\Big(\hat{f}_n(x) > \gamma_n(x), \quad \forall\, x \in \mathbb{R}\Big) \\ &\le P\Big(\hat{f}_n(x) > \eta_n, \quad \hat{g}_n(x) > \eta_n, \quad \forall\, x \in \mathcal{I}) \\ &\le P\Big(\hat{f}_n^b(x) > \eta_n - n^{-\epsilon/8}, \quad \hat{g}_n^b(x) > \eta_n - n^{-\epsilon/8}, \quad \forall\, x \in \mathcal{I}) + 2n^{-3} \\ &\le n^{-b+o(1)} \end{aligned}$$

(the last inequality follows by Theorem 1.4 for a threshold $\eta_n - n^{-\epsilon_n/8}$ when $x \in \mathcal{I} \cup \{x : x^{-1} \in \mathcal{I}\}$ and $-\infty$ otherwise). ☐

Lemma 5.2 *Suppose a_i are i.i.d. with $E(a_0) = 0$ and $E(a_0^2) = 1$. There exists $c < \infty$ such that for all $m = 2^{k+1}$ and k large enough,*

$$P\Big(f_m(x) > m^{-2}, \ \forall\, x \in [-1,1]\,, \qquad x g_m(x) \ge 0, \ \forall\, |x| \in [1-2^{-k}, 1]\Big) \ge m^{-c} \tag{5.11}$$

Proof: Define the intervals $\mathcal{J}_j = \{x : 1 - 2^{-j} \le |x| \le 1 - 2^{-j-1}\}$ for $j = 1, \dots, k-1$ and $\mathcal{J}_k = \{x : 1 - 2^{-k} \le |x| \le 1\}$. Throughout this proof, $l_j := 2^j$ for integer j, and complements are taken inside the interval $[-1, 1]$.

The proof of the lemma is based on decomposing f_m to a sum (over a number of terms logarithmic in m) of polynomials f^j, such that for each $x \in \mathcal{J}_j$, $f^j(x)$ is large while $f^i(x), i \ne j$ are not too large; at the same time, $g_m(x)$ is decomposed to a sum of polynomials all but the highest order of which are large and positive on $\mathcal{J}_k$, while the latter is not too negative on $\mathcal{J}_k$. Unfortunately, we need to introduce a few constants in order to define explicitly this decomposition.

Note first that for some $c_0 < \infty$ which does not depend on k,

$$(c_0 - 1)2^{j/2} x^{l_j} - \sum_{i=4, i\ne j}^{k} 2^{i/2} x^{l_i} \ge 0\,, \qquad \forall x \in \mathcal{J}_j, \ \ j = 4, \dots, k \tag{5.12}$$

Define $c_1 = c_0 + 1$. In Lemma 5.3 below, we define a constant $\theta_1 = \theta_1(c_1) > 0$. Define then $\theta = P(|N| \le 1)\theta_1/2 > 0$ where N is a standard Gaussian random variable of zero mean and unit variance. Since $E(a_0) = 0$, $E(a_0^2) = 1$, we can use Strassen's weak approximation theorem (see [Str] or [CS, Page 89]), to deduce the existence of independent standard normal random variables $\{b_i, i \ge 0\}$ such that, for all $j \ge j_0$,

$$P\left(\max_{0\le \ell \le 2^j}\left|\sum_{i=0}^{\ell} a_{2i} - \sum_{i=0}^{\ell} b_{2i}\right| + \max_{0\le \ell \le 2^j}\left|\sum_{i=0}^{\ell} a_{2i+1} - \sum_{i=0}^{\ell} b_{2i+1}\right| \ge 2^{j/2-3}\right) \le \theta. \tag{5.13}$$

Finally, since $Ea = 0$ and $Ea^2 = 1$ there exists $\alpha > 0$ such that $P(|a - \alpha| \le \delta) > 0$. Fixing such α, define $s > j_0$ such that

$$\frac{\alpha}{10} - \sum_{i=s}^{\infty} 2^{i/2} x^{l_i} \ge 0, \qquad \forall x \in \mathcal{J}_0 := \{x : |x| \le 1 - 2^{-s}\}. \tag{5.14}$$

Such an s always exists because the sum in (5.14) tends to 0 in s. Note that s does not depend on k and all estimates above are valid uniformly for all k large enough. We write $l := l_s$ and note that $\{\mathcal{J}_0, \mathcal{J}_s, \mathcal{J}_{s+1}, \ldots, \mathcal{J}_k\}$ form a partition of the interval $[-1,1]$. We keep s fixed throughout the rest of the proof.

As mentioned above, the proof of the lemma is based on decomposing f_m to a sum (over $k-s+2$ terms, i.e. a number of terms logarithmic in m) of polynomials $f^j, j = 0, s, s+1, \ldots, k$, while decomposing $g_m(x)$ to a similar sum of $k-s+2$ polynomials. Specifically, we write

$$f_m(x) = f^0(x) + \sum_{j=s}^{k} x^{l_j} f^j(x)$$

where $f^0 = f_{l_s}$ and

$$f^j(x) = \sum_{i=0}^{2^j-1} a_{i+l_j} x^i \ .$$

Similarly,

$$g_m(x) = x^{m-l_s} g^0(x) + \sum_{j=s}^{k} x^{m-l_{j+1}} g^j(x)$$

where $g^0 = g_{l_s}$ and

$$g^j(x) = \sum_{i=0}^{2^j-1} a_{i+l_j} x^{2^j-1-i} \ .$$

One checks that for k large enough, it holds that

$$m^{-2} \le \min\{\frac{\alpha}{10}, \inf_{x\in\mathcal{J}_j, j=s,\ldots,k} 2^{j/2} x^{l_j}\} \, . \tag{5.15}$$

Moreover, by (5.12), for all $k \ge s$,

$$(c_0-1)2^{j/2}x^{l_j} - \sum_{i=s, i\neq j}^{k} 2^{i/2} x^{l_i} \ge 0 \ , \qquad \forall x \in \mathcal{J}_j, \ \ j = s, \ldots, k \tag{5.16}$$

It follows that

$$\begin{aligned} \{f_m(x) > m^{-2}, \ \forall\, x \in [-1,1]\} \ &\supset \ \bigcap_{j=s}^{k} \{f^j(x) > c_0 2^{j/2}, \ \forall\, x \in \mathcal{J}_j, \ \ f^j(x) \ge -2^{j/2}, \ \forall\, x \in \mathcal{J}_j^c\} \\ &\cap \ \ \{f^0(x) > \frac{\alpha}{5}, \ \forall\, x \in \mathcal{J}_0, \quad f^0(x) \ge 0, \ \forall\, x \in \mathcal{J}_0^c\} \, . \end{aligned} \tag{5.17}$$

Note that for all $x \in [-1,1]$,

$$\{x g_m(x) \ge 0\} \supset \{x g^0(x) \ge -2^{k/2}\} \bigcap_{j=s}^{k} \{x g^j(x) \ge c_0 2^{j/2}\} \, . \tag{5.18}$$

The polynomial pairs (f^j, g^j), $j = 0, s, \ldots, k$ are mutually independent, with (f^0, g^0) having the same law as (f_l, g_l), while (f^j, g^j) has the same law as (f_{2^j}, g_{2^j}) for each $j \neq 0$. It thus follows from

(5.17) and (5.18) that

$$\begin{array}{ll} & P\Big(f_m(x) > m^{-2},\ \forall\ x\in[-1,1]\,, \qquad x g_m(x)\ge 0,\ \forall\ x\in \mathcal{J}_k\Big) \\ \geq & P\Big(f_l(x) > \dfrac{\alpha}{5},\ \ \forall\ x\in\mathcal{J}_0,\quad f_l(x)\ge 0,\ \ \forall\ x\in\mathcal{J}_0^c,\quad x g_l(x)\ge -2^{k/2},\ \ \forall\ x\in\mathcal{J}^k\Big) \\ \prod_{j=s}^k & P\Big(f_{2^j}(x) > c_0 2^{j/2},\ \ \forall\ x\in\mathcal{J}_j,\quad f_{2^j}(x)\ge -2^{j/2},\ \ \forall\ x\in\mathcal{J}_j^c,\quad x g_{2^j}(x)\ge c_0 2^{j/2},\ \ \forall\ x\in\mathcal{J}_k\Big) \\ := & \eta_{s,k}\displaystyle\prod_{j=s}^k q_j\,, \end{array} \tag{5.19}$$

where

$$\eta_{s,k} := P\Big(f_l(x) > \alpha/5,\ \ \forall\ x\in\mathcal{J}_0,\quad f_l(x)\ge 0,\ \ \forall\ x\in\mathcal{J}_0^c,\quad x g_l(x)\ge -2^{k/2},\ \ \forall\ x\in\mathcal{J}^k\Big)\,. \tag{5.20}$$

We first show that $\eta_{s,k}$ is uniformly (in k, for $k\ge 2\log_2(2\alpha l)$), bounded away from zero, and then provide a uniform (in k) bound (independent of j) on q_j. Toward the first goal, let $Q_s(x) := \alpha(1+x+\cdots+x^{l_s-1})$, noting that $Q_s(x)$ is monotone increasing on $[-1,1]$, with $Q_s(-(1-2^{-s}))\ge\alpha/4$, implying that $Q_s(x)\ge\alpha/4$ for all $x\in\mathcal{J}_0$. Thus, for each $s\ge 1$ there exists $\delta_s\in(0,\alpha)$ such that $f_{l_s}(x) > \alpha/5$ whenever $x\in\mathcal{J}_0$ and $|a_i-\alpha|\le\delta_s$ for $i=0,\dots,l_s-1$. Further taking $a_{2i}\ge a_{2i+1}\ge 0$ for $i=0,\dots,2^{s-1}-1$ guarantees that $f_{l_s}(x)\ge 0$ for all $x\in[-1,1]$. Considering only such $\{a_i\}$, we also have that $|x g_l(x)|\le 2\alpha l$, and hence, combining the above and using $2^{k/2}\ge 2\alpha l$, we have that

$$\liminf_{k\to\infty}\eta_{s,k} > 0\,. \tag{5.21}$$

To estimate q_j, we note that $x g^b_{2^j} = x^2 g^b_{2^j-1} + x b_{2^j-1}$. Thus, combining (5.1), (5.3) and (5.13), it follows that for all $j\in\{s,\dots,k\}$,

$$q_j \ge P(|b_{2^j-1}|\le 1)q^b_j - \theta \tag{5.22}$$

where

$$q^b_j := P\Big(f^b_{2^j-1}(x)\ge c_1 2^{j/2},\ \forall\, x\in\mathcal{J}_j,\ \ f^b_{2^j-1}(x)\ge -2^{(j-1)/2},\ \forall\, x\in\mathcal{J}_j^c,\ \ g^b_{2^j-1}(x)\ge c_1 2^{j/2},\ \forall\, x\in\mathcal{J}_k\Big)\,.$$

for say, $c_1 = c_0+1$. Slepian's lemma thus yields that for all $k\ge j\ge s$,

$$q^b_j \ge P\Big(f^b_{2^j-1}(x)\ge c_1 2^{j/2},\ \forall x\in[1-2^{-j},1]\Big)^4 P\Big(f^b_{2^j-1}(x)\ge -2^{(j-1)/2},\ \forall x\in[0,1-2^{-j}]\Big)^2 := \bar{q}^b_j.$$

Note that $\bar{q}^b_j$ does not depend on k, and in fact it depends on c_1 and j only. The following lemma provides estimates on $\bar{q}^b_j$ while defining the constant θ_1:

Lemma 5.3 *There exists a constant $\theta_1>0$ such that for all $j\ge 4$,*

$$\bar{q}^b_j \ge \theta_1\,.$$

Applying (5.22) using $\theta=\frac{1}{2}P(|b|\le 1)\theta_1$ then leads to $q^j\ge\theta$ for all $j\ge s$. In view of (5.21) and (5.19) this proves (5.11). $\square$

Proof of Lemma 5.3: Note that $\sigma_{2^j-1}(x)\ge 2^{j/2-1}$ when $x\in[1-2^{-j},1]$, hence by Lemma 2.4, for some $\xi_1>0$ and all j large enough,

$$P\Big(f^b_{2^j-1}(x)\ge c_1 2^{j/2},\ \ \forall x\in[1-2^{-j},1]\Big)\ge P\Big(\hat{f}^b_{2^j-1}(x)\ge 2c_1,\ \ \forall x\in[1-1/(2^j-1),1]\Big)\ge\xi_1\,.$$

Note that $\sigma_{2^j-1}(x) \le 1/\sqrt{1-x} \le 2^{i/2}$ for $x \in [0, 1-2^{-i}]$. Hence, by Slepian's lemma and (3.3), we have that for the Ornstein-Uhlenbeck process X_t of (2.14),

$$
\begin{aligned}
&P\Big(f^b_{2^j-1}(x) \ge -2^{(j-1)/2}, \forall x \in [0, 1-2^{-j}]\Big) \\
&\ge \prod_{i=0}^{j-1} P\Big(f^b_{2^j-1}(x) \ge -2^{(j-1)/2},\ \forall x \in [1-2^{-i}, 1-2^{-i-1}]\Big) \\
&\ge \prod_{i=0}^{j-1} P\Big(X_t \ge -2^{(j-i)/2-1},\ \forall t \in [i \ln 2, (i+1)\ln 2]\Big) \\
&= \prod_{l=1}^{j} P\Big(\inf_{0\le t\le \ln 2} X_t \ge -2^{l/2-1}\Big) \\
&\ge \prod_{l=1}^{\infty} \Big[1 - P(\sup_{0\le t\le \ln 2} X_t \ge 2^{l/2-1})\Big] =: \xi_2
\end{aligned}
$$

and $\xi_2 > 0$ since $E(\sup_{t\in[0,\ln 2]} X_t) < \infty$. This completes the proof. □

6 Proof of Theorem 1.1

Part a) of Theorem 1.1 is a direct consequence of Theorem 1.3 with $\gamma_n = 0$. Thus, it only remains to prove part b). Fixing $\mu \ne 0$, it is easy to see that

$$P_n^\mu = P(\hat{f}_n(x) \ne -\mu\kappa_n(x), \quad \forall x \in \mathbb{R}),$$

where the nonrandom $\kappa_n(x) = (\sum_{i=0}^{n-1} x^i)(\sum_{i=0}^{n-1} x^{2i})^{-1/2}$ are strictly positive and $\hat{f}_n(x)$ are the normalized polynomials that correspond to a_i of zero mean. With $\widetilde{P}_n$ for the value of P_n when coefficients $\{-a_i\}$ are used instead of $\{a_i\}$, it is easy to see that

$$P_n^\mu = P_{n,-\mu\kappa_n} + \widetilde{P}_{n,\mu\kappa_n}.$$

Consequently, we may and shall assume without loss of generality that $\mu > 0$, proving only that $P_{n,-\mu\kappa_n} = n^{-b/2+o(1)}$. Observe that $\kappa_n(1) = \sqrt{n}$, $\kappa_n(-1) = 1/\sqrt{n}$ and if $|x| \ne 1$ then,

$$\kappa_n(x) = \kappa_n(x^{-1}) = \kappa_n(-x)^{-1} = \left[\frac{(1+x)(1-x^n)}{(1-x)(1+x^n)}\right]^{1/2}. \tag{6.1}$$

Moreover, there exists $c = c(\mu) > 0$ such that for all n large enough,

$$\mu\kappa_n(x) \ge n^{-\epsilon/8} + \frac{c}{\sqrt{1-x}} \qquad \forall x \in [0, 1-n^{-1}]. \tag{6.2}$$

For an upper bound on $P_{n,-\mu\kappa_n}$ let $\mathcal{I}_- = [-1, -1+0.5n^{-\epsilon}]$ be the subset of $\mathcal{I}$ of Section 5 near the point -1 and $V_- = \mathcal{I}_3 \cup \mathcal{I}_4$ be the (corresponding) subset of V of Section 4. It is easy to check that

$$\sup\{\kappa_n(x) : x \in \mathcal{I}_-\} \le c_1 n^{-\epsilon/2}$$

for some $c_1 < \infty$ and all n. Hence, applying the arguments of Section 5 followed by those of Section 4 with $\mathcal{I}_-$ replacing $\mathcal{I}$ and V_- replacing V, respectively, results in the stated upper bound $P_{n,-\mu\kappa_n} \le n^{-b/2+o(1)}$. Turning to prove the corresponding lower bound on $P_{n,-\mu\kappa_n}$, let $\mathcal{I}_+ = [1-0.5n^{-\epsilon}, 1]$

denote the subset of $\mathcal{I}$ near the point $+1$. It follows from (6.1) that ρ_n of (5.6) is zero for $\gamma_n = -\mu\kappa_n$, allowing for the use of $r_n = n^{-1}$ and $\xi_n(x) = -\mu x^m \sigma_{n-2m}(x)\kappa_n(x) \leq 0$ in (5.8). We then deal with the terms Q_2, Q_3 and Q_4 as in Section 5. For the dominant term Q_1, instead of (5.9) we have in view of (6.2) that

$$\begin{aligned} Q_1 &\geq P\Big(\hat{f}^b_{n'}(x) > n^{-\epsilon/8}, \quad \hat{g}^b_{n'}(x) > n^{-\epsilon/8}, \quad \forall\, x \in [-1,0] \cup [1-\frac{1}{n'},1], \\ &\qquad \hat{f}^b_{n'}(x) > -\frac{c}{\sqrt{1-x}}, \quad \hat{g}^b_{n'}(x) > -\frac{c}{\sqrt{1-x}}, \quad \forall\, x \in [0, 1-\frac{1}{n'}]\Big) - 2n^{-3} \\ &= \widetilde{Q}_1 - 2n^{-3}, \end{aligned} \tag{6.3}$$

where $n' := n - 2m = n(1+o(1))$. By Slepian's lemma, similarly to (3.1) we see that

$$\begin{aligned} \widetilde{Q}_1 &\geq P\Big(\hat{f}^b_{n'}(x) > -\frac{c}{\sqrt{1-x}} \quad \forall\, x \in [0, 1-\frac{1}{n'}]\Big)^2 \\ &\qquad P\Big(\hat{f}^b_{n'}(x) > n^{-\epsilon/8} \quad \forall\, x \in [0,1]\Big)^2 P\Big(\hat{f}^b_{n'}(x) > 1 \quad \forall\, x \in [1-\frac{1}{n'},1]\Big)^2 \\ &:= (A_{n'})^2 (B_{n'})^2 (C_{n'})^2 \,. \end{aligned} \tag{6.4}$$

The sequence $C_{n'}$ is bounded away from zero by Lemma 2.4. Moreover, it is shown in Section 3 that $B_n \geq n^{-b/4+o(1)}$. In view of (6.3) and (6.4), it thus suffices to show that the sequence A_n is bounded below by some $A_\infty > 0$ in order to conclude that $P^\mu_n \geq P_{n,-\mu\kappa_n} \geq n^{-b/2+o(1)}$ and complete the proof of part b) of Theorem 1.1. To this end, recall that the function $\sqrt{(1-x\vee y)/(1-x\wedge y)}$ in the right side of (3.3) is the covariance of the process $W_{1-x}/\sqrt{1-x}$. Consequently, we have by (3.3) and Slepian's lemma that

$$A_n \geq P(W_{1-x} > -c, \forall x \in [0, 1-n^{-1}]) \geq P(\inf_{0\leq x\leq 1} W_x > -c) = A_\infty > 0\,,$$

as needed. □

7 Upper bound for Theorem 1.2

Fixing small $\delta > 0$ and integers $k_n = o(\log n/\log\log n)$, it suffices for the upper bound in Theorem 1.2 to provide an $n^{-(1-2\delta)b+o(1)}$ upper bound on the probability $q_{n,k}$ that $f(x) = \hat{f}_{n+1}(x)$ has at most $k = k_n$ zeros in the set $V = \cup_{i=1}^4 \mathcal{I}_i$ of Section 4. To this end, let $x = \theta_i(1-e^{-t})$ within $\mathcal{I}_i$, $i = 1,2,3,4$, where θ_i is defined in Section 3. With $T = \log n$ cut the range $t \in [\delta T, (1-\delta)T]$ for each $\mathcal{I}_i$ to $(1-2\delta)T$ unit length intervals, denoting by $J_{(i-1)T+1}, \ldots, J_{iT}$ the corresponding image in $\mathcal{I}_i$. If $f(x)$ has ℓ zeros in some $\mathcal{I}_i$, then there must exist $j_1, \ldots, j_\ell$ such that $f(x)$ has a constant sign $s \in \{-1,1\}$ on each of the "long" subintervals obtained by deleting $J_{(i-1)T+j_1}, \ldots, J_{(i-1)T+j_\ell}$ from $\mathcal{I}_i$. We partition the event that $f(x)$ has at most k zeros in V according to the possible vector $\mathbf{j} = (j_1, \ldots, j_k)$ of "crossing indices" among the $4T$ intervals $\{J_1, \ldots, J_{4T}\}$ and the possible signs $s_m \in \{-1,1\}$ of $f(x)$ on the resulting long subintervals L_m, $m = 1, \ldots, k+4$ within V. Let

$$q_{n,\mathbf{s},\mathbf{j}} = P\Big(\min_{m=1}^{k+4} \inf_{x\in L_m} s_m f(x) > 0\Big)\,,$$

for $\mathbf{s} = (s_1, \ldots, s_{k+4})$. Since

$$q_{n,k} \leq \sum_{\mathbf{j}} \sum_{\mathbf{s}} q_{n,\mathbf{s},\mathbf{j}}\,,$$

and the number of choices of $\mathbf{j}$ and $\mathbf{s}$ is at most $2^{k+4}(4T)^k = n^{o(1)}$, it suffices to show that

$$\max_{\mathbf{s},\mathbf{j}} q_{n,\mathbf{s},\mathbf{j}} \leq n^{-(1-2\delta)b+o(1)} . \tag{7.1}$$

Applying the coupling of (f_k, g_k) and (f_k^b, g_k^b) as provided in (5.4) and (5.5) for $t = n^{\delta/4}$ and $p = 16/\delta$, we see that for all $\mathbf{j}$ and $\mathbf{s}$,

$$q_{n,\mathbf{s},\mathbf{j}} \leq P\Big(\min_{m=1}^{k+4} \inf_{x \in L_m} s_m \hat{f}_{n+1}^b(x) > -n^{-\delta/8}\Big) + cn^{-3} := q_{n,\mathbf{s},\mathbf{j}}^b + cn^{-3} , \tag{7.2}$$

where $c < \infty$ depends on δ but is independent of $\mathbf{j}$, $\mathbf{s}$ and n. Thus, the proof reduces to the Gaussian case.

Suppose first that n is even. The covariance function of $s_m \hat{f}_{n+1}^b(x)$ on $V' = \cup_m L_m$ is then given by $s_m s_l c_{n+1}(x, y)$ for $x \in L_m$ and $y \in L_l$. Since $c_{n+1}(x, y) \geq 0$ for all x, y, it follows by Slepian's lemma that per choice of $\mathbf{j}$, the probability $q_{n,\mathbf{s},\mathbf{j}}^b$ is maximal when $s_m = 1$ for all m. In case n is odd, note that $f_{n+1}^b(x) = \sigma_n(x) \hat{f}_n^b(x) + b_n x^n$ and

$$|x|^n \leq 2n^{-\delta/2} \sigma_n(x) \qquad \forall\, x \in V$$

Consequently, for all $\mathbf{j}$ and $\mathbf{s}$,

$$q_{n,\mathbf{s},\mathbf{j}}^b \leq P(|b_n| \geq n^{\delta/4}) + P\Big(\min_{m=1}^{k+4} \inf_{x \in L_m} s_m \hat{f}_n^b(x) > -2n^{-\delta/8}\Big) . \tag{7.3}$$

With $n-1$ even, continuing as before, we see that the right-most term in (7.3) is maximal when $s_m = 1$ for all m. In conclusion, it suffices to consider

$$q_{n,\mathbf{j}}^b = P\Big(\inf_{x \in V'} \hat{f}_{n+1}^b(x) > -2n^{-\delta/8}\Big) ,$$

for n even. Applying the arguments of (4.2), (4.4) and (4.5) with $\gamma_n(x) \equiv -n^{-\delta/8}$ on the subset V' of V, we find that

$$q_{n,\mathbf{j}}^b \leq e^{-n^{\delta/4}/2} + \prod_{i=1}^{4} P\Big(\sup_{t \in \mathcal{T}_i} Y_t \leq \varepsilon_{T'}\Big) , \tag{7.4}$$

where Y_t is the stationary Gaussian process of Lemma 2.2 and for $i = 1, \ldots, 4$ the set $\mathcal{T}_i \subset [\delta T, (1-\delta)T]$ is the image of $V' \cap \mathcal{I}_i$ under the transformation $t = -\log(1-\theta_i(x))$. Since $\tau \mapsto R_y(\tau)$ is monotonically decreasing on $[0, \infty)$, it follows by Slepian's lemma that $P(\sup_{t \in \mathcal{T}_i} Y_t \leq \varepsilon_{T'})$ is maximal per fixed size of $\mathcal{T}_i$ when the latter set is an interval, that is, when the J_{j_l} are all at one end of $\mathcal{I}_i$ for each i (easiest to see this by considering first J_{j_1} only, then J_{j_2} etc.). In this case each interval $\mathcal{T}_i$ has at least the length $(1-2\delta)T - k$, so the upper bound of (7.1) follows from (2.5) and (7.2)–(7.4). $\square$

8 Proof of Theorem 1.2

In view of the upper bound $q_{n,k} \leq n^{-b+o(1)}$ of Section 7, it suffices to show that $p_{n,k} \geq n^{-b-o(1)}$, in order to complete the proof of Theorem 1.2.

Since a_i are of zero mean and positive variance, the support of their law must intersect both $(0, \infty)$ and $(-\infty, 0)$. Consequently, there exist $\beta < 0 < \alpha$ such that $P(|a_i - \alpha| < \epsilon) > 0$ and $P(|a_i - \beta| < \epsilon) > 0$ for all $\epsilon > 0$. Replacing $\{a_i\}$ by $\{-a_i\}$ does not affect the number of zeros of

$f_{n+1}(x)$. Hence, we may and shall assume without loss of generality that $|\alpha| \geq |\beta|$. Let $s \geq 4$ be an even integer such that $\alpha + (s-1)\beta < 0$. Define

$$Q(x) = \beta x^{s-1} + \sum_{i=0}^{s-2} \alpha x^i, \qquad R(x) = \alpha + \sum_{i=1}^{s-1} \beta x^i,$$

and note that

$$Q(x) > 0 \quad \forall |x| \leq 1, \qquad R(1) < 0 < R(-1)\,. \tag{8.1}$$

8.1 Proof for k and n even

Suppose that k and n are even. After k, s, α, β are fixed, we shall choose $\delta > 0$ sufficiently small, then a large enough integer $r = r(\delta)$, followed by a small enough positive $\epsilon = \epsilon(\delta, r)$, all of which are independent of n. Let r_i denote the multiple of s nearest to r^i, for $i = 1, \ldots, k$ and $\rho_n := \max\{5/p_n, (\log n)^{-1/2}\}$ for $p_n \uparrow \infty$ such that $E|a_i|^{p_n} \leq n$ (these choices are slightly different from the ones made in Section 5). Let $m = m_n$ be the multiple of s nearest to $2r_k\rho_n \log n / |\log(1-\delta)|$ and define the polynomials

$$\begin{aligned}
B(x) = \sum_{i=0}^{m-1} b_i x^i \quad &:= \quad (1 + x^s + x^{2s} + \ldots + x^{r_1 - s})Q(x) + (x^{r_1} + x^{r_1+s} + \ldots + x^{r_2-s})R(x) \\
&+ \quad (x^{r_2} + x^{r_2+s} + \ldots + x^{r_3-s})Q(x) + (x^{r_3} + x^{r_3+s} + \ldots + x^{r_4-s})R(x) \\
&+ \quad \ldots + (x^{r_k} + x^{r_k+s} + \ldots + x^{m-s})Q(x), \\
C(X) = \sum_{i=0}^{m-1} c_i X^i \quad &:= \quad (1 + X^s + X^{2s} + \ldots + X^{m-s})Q(X).
\end{aligned}$$

Each coefficient b_i of $B(x)$ equals either α or β. The same holds for each coefficient c_i of $C(X)$. Let $\mathcal{A}_n$ denote the event that the following hold:

A1 $\quad |a_i - b_i| < \epsilon$ for $i = 0, 1, \ldots, m-1$

A2 $\quad |a_{n-i} - c_i| < \epsilon$ for $i = 0, 1, \ldots, m-1$

A3 $\quad a_m + a_{m+1}x + \ldots + a_{n-m}x^{n-2m} > n^{-1/4}\sigma_{n-2m+1}(x)$ for all $x \in \mathbb{R}$.

A4 $\quad |a_i| < n^{\rho_n}$ for $i = 0, 1, \ldots, n$.

Most of our work shall be to show that the polynomial $B(x)$ has the required behavior in terms of zeros for $|x| \leq (1-\delta)^{1/r_k}$. Conditions **A1** and **A4** ensure that $f(x)$ is close enough to $B(x)$ on this interval so as to have there exactly k simple zeros. The condition **A3** precludes additional zeros of $f(x)$ near ± 1. Moreover, with **A2** and the positivity of $C(X)$ for $|X| \leq 1$, we conclude that $f(x) > 0$ when $|x| > 1$.

The stated lower bound on $p_{n,k}$ is an immediate consequence of the following two lemmas.

Lemma 8.1 *For any fixed $\delta > 0$, $\epsilon > 0$ and an integer r, the probability of the event $\mathcal{A}_n$ is at least $n^{-b-o(1)}$ for even $n \to \infty$.*

Lemma 8.2 *Suppose the even integer k is fixed. There exist small enough $\delta > 0$, large enough $r = r(\delta)$ and a small enough $\epsilon = \epsilon(\delta, r)$ positive, such that for all sufficiently large even n, any polynomial $f(x) = \sum_{i=0}^{n} a_i x^i$ whose coefficients are in $\mathcal{A}_n$ has exactly k real zeros, each of which is a simple zero.*

Proof of Lemma 8.1: Since all coefficients of $B(x)$ and $C(x)$ belong to $\{\alpha, \beta\}$, our choice of α and β implies that each coefficient condition in **A1** and **A2** is satisfied with probability at least c for some $c > 0$ depending only on ϵ. The probability that condition **A3** holds is $P_{n',\gamma_{n'}}$ of Theorem 1.3 for $n' = n - 2m + 1$ odd and $\gamma_{n'}(x) \equiv n^{-1/4}$ that satisfy the assumptions of this theorem. Consequently, condition **A3** holds with probability of at least $(n')^{-b-o(1)} = n^{-b-o(1)}$. Since conditions **A1**, **A2** and **A3** are independent, the probability that all of them hold is at least

$$c^{2m}(n')^{-b+o(1)} = n^{-b-o(1)}. \tag{8.2}$$

(Recall that $\rho_n \to 0$, hence also $m_n/\log n \to 0$.) By Markov's inequality and the choice of ρ_n, the probability that condition **A4** fails for a given i is at most n^{-4}. Hence the probability that this condition fails for any i in the range $0 \le i \le n$ is at most $O(n^{-3})$. Since $b \le 2$, imposing condition **A4** does not affect the $n^{-b-o(1)}$ lower bound of (8.2). $\square$

Proof of Lemma 8.2: The proof of the lemma is divided in three steps.
Step 1: For $\delta > 0$ sufficiently small, $r > (\log \delta)/(\log(1-\delta))$ sufficiently large, $\epsilon > 0$ sufficiently small and all large even integers n, each polynomial $f(x)$ with coefficients in $\mathcal{A}_n$ has exactly k simple zeros in $[0, 1]$.
Step 2: Under same conditions on the parameters, $f(x) > 0$ on $[-1, 0]$.
Step 3: $g(X) = X^n f(X^{-1}) > 0$ on $(-1, 1)$.
Step 1. Fixing $f(x)$ as above, observe that the zeros of $f(x)$ in $(0, 1)$ are the same as those of $F(x) := (1 - x^s)f(x)$, so it suffices to prove that

- $F(x) > 0$ for $x \in [0, \delta^{1/r_1}]$
- $F'(x) < 0$ for $x \in [\delta^{1/r_1}, (1-\delta)^{1/r_1}]$
- $F(x) < 0$ for $x \in [(1-\delta)^{1/r_1}, \delta^{1/r_2}]$
- $F'(x) > 0$ for $x \in [\delta^{1/r_2}, (1-\delta)^{1/r_2}]$
- $F(x) > 0$ for $x \in [(1-\delta)^{1/r_2}, \delta^{1/r_3}]$
- $F'(x) < 0$ for $x \in [\delta^{1/r_3}, (1-\delta)^{1/r_4}]$
  $\vdots$
- $F'(x) > 0$ for $x \in [\delta^{1/r_k}, (1-\delta)^{1/r_k}]$
- $F(x) > 0$ for $x \in [(1-\delta)^{1/r_k}, 2^{-1/m})$
- $f(x) > 0$ for $x \in [2^{-1/m}, 1]$

Indeed, the sign changes in $F(x)$ force at least one real zero in each of the k gaps *between* the intervals on which F is guaranteed positive or negative, and the monotonicity of $F'(x)$ on these gaps guarantees that each of them contains exactly one zero and that the zero is simple. Note also that $m = m_n \to \infty$, so per choice of $\delta > 0$ all the intervals of x above are nonempty as soon as r is large enough.

Recall that our choice of $m = m_n$ is such that for any $l < \infty$,

$$m^l n^\rho (1-\delta)^{m/r_k} \to 0\ . \tag{8.3}$$

Consequently, by conditions **A1** and **A4** of $\mathcal{A}_n$, there exists $c(r,\delta)$ finite, such that for all $\epsilon > 0$, n large enough and $|x| \le (1-\delta)^{1/r_k}$,

$$\begin{aligned} |F(x) - (1-x^s)B(x)| &\le (1-x^s)\left[\epsilon(1+|x|+\ldots+|x|^{m-1}) + n^\rho(|x|^m+\ldots+|x|^n)\right] \\ &\le (1-|x|)^{-1}(\epsilon + n^\rho x^m) \le c(r,\delta)\epsilon. \end{aligned} \tag{8.4}$$

Fix M such that $|Q(x)| \le M$ and $|R(x)| \le M$ for all $x \in [0,1]$. By definition of $B(x)$,

$$(1-x^s)B(x) = Q(x) + \left[\sum_{\ell=1}^{k}(-1)^\ell(Q(x)-R(x))x^{r_\ell}\right] - Q(x)x^m. \tag{8.5}$$

Suppose $x \in [0, \delta^{1/r_1}]$. Then, each x^{r_ℓ} and x^m is at most δ, so

$$(1-x^s)B(x) \ge Q(x) - (2k+1)M\delta \,.$$

Therefore, for all δ sufficiently small, the positivity of $Q(x)$ on $[0,1]$ (see (8.1)) implies that $(1-x^s)B(x) \ge \eta$ for some $\eta > 0$ independent of n, and all $x \in [0,\delta^{1/r_1}]$. For $\epsilon > 0$ small enough, this in turn implies the positivity of $F(x)$ on this interval (see (8.4)).

Suppose $x \in [(1-\delta)^{1/r_j}, \delta^{1/r_{j+1}}]$ for some $j \in \{1,2,\ldots,k-1\}$. Then, $x^m \le x^{r_\ell} \le \delta$ for all $\ell > j$ and $x^{r_\ell} \in [1-\delta, 1]$, for all $\ell \le j$. In view of the identity

$$Q(x) + \sum_{\ell=1}^{j}(-1)^\ell(Q(x)-R(x)) = Q(x)1_{j \textit{ is even}} + R(x)1_{j \textit{ is odd}} \tag{8.6}$$

and (8.5), it follows that for all x as above,

$$|(1-x^s)B(x) - [Q(x)1_{j \textit{ is even}} + R(x)1_{j \textit{ is odd}}]| \le (2k+1)M\delta.$$

For δ small enough, the error $(2k+1)M\delta$ is at most $\min\{Q(1), -R(1)\}/3$. Once δ is chosen, taking r sufficiently large guarantees that $Q(x) \ge Q(1)/2$ and $R(x) \le R(1)/2$ for all $x \in [(1-\delta)^{1/r_1}, 1]$. Since $Q(1)$ is positive and $R(1)$ is negative (see (8.1)), we conclude that there exists $\eta > 0$ independent of n such that $(-1)^j(1-x^s)B(x) \ge \eta$ for all n large enough and all $x \in [(1-\delta)^{1/r_j}, \delta^{1/r_{j+1}}]$, $j = 1,\ldots,k-1$. In view of (8.4), for all $\epsilon > 0$ small enough $(-1)^jF(x)$ is then positive throughout the interval $x \in [(1-\delta)^{1/r_j}, \delta^{1/r_{j+1}}]$, as needed.

Suppose $x \in [(1-\delta)^{1/r_k}, 2^{-1/m}]$. Then, $x^{r_\ell} \in [1-\delta, 1]$ for all $\ell \le k$ and $x^m \le 1/2$. With k even, it follows from (8.5) and (8.6) that

$$(1-x^s)B(x) \ge \frac{1}{2}Q(x) - 2kM\delta \,.$$

So, when $\delta > 0$ is small enough, then for some $\eta > 0$ independent of n, it holds that $(1-x^s)B(x) \ge \eta$ for all n large enough and all $x \in [(1-\delta)^{1/r_k}, 2^{-1/m}]$. Recall that

$$(1-x^s)x^m(a_m + a_{m+1}x + \ldots + a_{n-m}x^{n-2m}) \ge 0 \tag{8.7}$$

by condition **A3**. So, while $F(x) - (1-x^s)B(x)$ is no longer negligible as in (8.4), the positivity of the expression in (8.7) results in

$$F(x) - (1-x^s)B(x) \ge -(1-x^s)(\epsilon(1+|x|+\ldots+|x|^{m-1}) + m(\alpha+\epsilon)|x|^{n-m}) \ge -c(\delta,r)\epsilon$$

for some finite $c(\delta, r)$, all $\epsilon > 0$ and large enough n. (This is because $m_n = o(\log n)$, so $m_n x^{n-m_n} \to 0$ as $n \to \infty$, uniformly on $x \in [(1-\delta)^{1/r_k}, 2^{-1/m}]$). Consequently, when $\epsilon > 0$ is small enough, the uniform positivity of $(1-x^s)B(x) \geq \eta > 0$ results in the positivity of $F(x)$ for $x \in [(1-\delta)^{1/r_k}, 2^{-1/m}]$.

Suppose $x \in [2^{-1/m}, \delta^{1/n}]$. Using the decomposition $f = f^L + f^M + f^H$ as in (5.7), note that by condition **A1**

$$\begin{aligned} f^L(x) \geq B(x) - \epsilon \sum_{i=0}^{m-1} |x|^i &\geq (x^{r_k} + x^{r_k+s} + \ldots + x^{m-s})Q(x) - r_k M - \epsilon m \\ &\geq \left(\frac{m - r_k}{2s}\right) Q(x) - r_k M - \epsilon m. \end{aligned}$$

Note that $f^M(x) \geq 0$ by condition **A3** and

$$|f^H(x)| \leq m(\alpha + \epsilon)x^{n-m} \leq m(2\alpha)\delta^{1-m/n} \leq 2m\alpha\delta^{1/2}$$

by condition **A2**. Since $m_n \to \infty$ while $m_n/n \to 0$, and $Q(x)$ is strictly positive, we see by combining the above that if δ and ϵ are small enough then for all n large enough the "main" term $m_n Q(x)/(2s)$ dominates, so $f(x) > 0$ for all $x \in [2^{-1/m}, \delta^{1/n}]$.

Suppose $x \in [\delta^{1/n}, 1]$. In this case, by condition **A3**,

$$f^M(x) \geq x^m n^{-1/4} \sigma_{n-2m+1}(x) \geq \delta^2 n^{-1/4} \sqrt{n - 2m + 1} > n^{1/8} \tag{8.8}$$

as $n \to \infty$. Condition **A1** implies that $|f^L(x)| \leq (\alpha + \epsilon)m$, whereas condition **A2** implies that $|f^H(x)| \leq (\alpha + \epsilon)m$. Since $m = o(\log n)$, we conclude that $f(x) > 0$ for large n and all $x \in [\delta^{1/n}, 1]$.

We turn to deal with the sign of $F'(x)$ in the gaps $[\delta^{1/r_j}, (1-\delta)^{1/r_j}]$ for $j = 1, \ldots, k$. To this end, first note that

$$F'(x) = \frac{d}{dx}\left[(1 - x^s)B(x)\right] + e(x)$$

where by conditions **A1** and **A4**, there exists $c(\delta, r)$ finite, such that for all $\epsilon > 0$, n large enough and $x \in [0, (1-\delta)^{1/r_k}]$,

$$\begin{aligned} |e(x)| &\leq | - s x^{s-1}| \left[\epsilon(1 + x + \ldots + x^{m-1}) + n^\rho (x^m + x^{m+1} + \ldots + x^n)\right] \\ &\quad + (1 - x^s)\left[\epsilon(1 + 2x + \ldots + (m-1)x^{m-2}) + n^\rho (m x^{m-1} + (m+1)x^m + \ldots + n x^{n-1})\right] \\ &\leq s(1-x)^{-1}(\epsilon + n^\rho x^m) + (1-x)^{-2}(\epsilon + n^\rho m x^{m-1}) \leq c(\delta, r)\epsilon \end{aligned} \tag{8.9}$$

(see (8.3) and (8.4)). Next, using (8.5), we obtain

$$\frac{d}{dx}\left[(1 - x^s)B(x)\right] = Q'(x) + \sum_{\ell=1}^{k} (-1)^\ell \left[(Q'(x) - R'(x))x^{r_\ell} + (Q(x) - R(x)) r_\ell x^{r_\ell - 1}\right] - o(1). \tag{8.10}$$

(The $o(1)$ denotes two terms involving x^m, which by (8.3) converge to 0 uniformly on $x \in [0, (1-\delta)^{1/r_k}]$.) The sum of the terms involving $Q'(x)$ or $R'(x)$ in (8.10) is at most $(2k+2)M'$, where M' is such that $|Q'(x)| \leq M'$ and $|R'(x)| \leq M'$ for all $x \in [0, 1]$. Per fixed $\delta > 0$, if r is sufficiently large then $Q(x) - R(x) \geq \eta$ for some $\eta > 0$ and all $x \in [\delta^{1/r_1}, 1]$ (see (8.1)). We claim that if $x \in [\delta^{1/r_j}, (1-\delta)^{1/r_j}]$ for some $j \in \{1, 2, \ldots, k\}$ then the term $h_j := (-1)^j (Q(x) - R(x)) r_j x^{r_j - 1}$ dominates the right hand side of (8.10) for all r large enough. Indeed, $|h_j| \geq \eta\delta r_j$ for all $x \in [\delta^{1/r_j}, (1-\delta)^{1/r_j}]$, whereas for such x we have that $|h_\ell| \le 2Mr_{j-1}$ when $\ell < j$ and $|h_\ell| \le 2Mr_k(1-\delta)^{(r_{j+1}-1)/r_j}$ when $\ell > j$. Since $r_\ell = r^\ell(1+o(1))$, combining the above we see that for all large enough r,

$$(-1)^j \frac{d}{dx}[(1-x^s)B(x)] \ge \frac{\eta}{2}\delta r^j - 3Mk(r^{j-1} + r^k(1-\delta)^{\sqrt{r}}) - (2k+2)M' - o(1) \ge \frac{\eta}{3}\delta r^j .$$

By (8.9) we then get that for small enough $\epsilon > 0$, $(-1)^j F'(x)$ also is positive in the j-th gap.

This completes Step 1.

Step 2. As before, define $F(x) := (1-x^s)f(x)$. The proof that $F(x) > 0$ on $[-\delta^{1/r_1}, 0]$ is the same as the proof for $[0, \delta^{1/r_1}]$, now using the positivity of $Q(x)$ on $[-1, 0]$. For each $j \in \{1, 2, \ldots, k-1\}$, the analysis for $[-\delta^{1/r_{j+1}}, -(1-\delta)^{1/r_j}]$ is the same as that for $[(1-\delta)^{1/r_j}, \delta^{1/r_{j+1}}]$, the only difference is that $Q(x)$ and $R(x)$ are both positive near -1 (whereas they have opposite signs near 1), so the result is that $F(x) > 0$ on these intervals, independent of the parity of j. The analyses for $[-2^{-1/m}, -(1-\delta)^{1/r_k}]$ and for $[-1, -2^{-1/m}]$ are the same as for the symmetric intervals on the positive side.

To complete the proof that $f(x) > 0$ on $[-1, 0]$, it remains to show that $F(x) > 0$ on each gap $[-(1-\delta)^{1/r_j}, -\delta^{1/r_j}]$ for $j = 1, \ldots, k$. By (8.4), it suffices to show that on such an interval $(1-x^s)B(x) \ge \eta$ for some $\eta > 0$, independent of ϵ and n. On the j-th such interval, $x^m \le x^{r_\ell} \le \delta$ for all $\ell > j$, whereas if r is sufficiently large, then $1 \ge x^{r_\ell} \ge \delta^{r_\ell/r_j} \ge (1-\delta)$ for all $\ell < j$. Hence, it follows from (8.5) and (8.6) that

$$|(1-x^s)B(x) - [t(x)Q(x) + (1-t(x))R(x)]| \le (2k-1)M\delta , \tag{8.11}$$

where $t(x) = 1 - x^{r_j}$ for j even, and $t(x) = x^{r_j}$ otherwise. Let $\eta = \min\{Q(-1), R(-1)\}/4 > 0$, and take δ small enough that $(2k-1)M\delta < \eta$. Since $t(x) \in [0, 1]$, if r is large enough that $\min\{Q(x), R(x)\} > 2\eta$ for all $x \in [-1, -\delta^{1/r_1}]$, then (8.11) implies that $(1-x^s)B(x) \ge \eta$ for all $x \in [-(1-\delta)^{1/r_j}, -\delta^{1/r_j}]$, $j = 1, \ldots, k$. The positivity of $F(x)$ for small ϵ and large n follows (by (8.4)).

Step 3. To complete the proof of Lemma 8.2, it suffices to show that $g(X) := X^n f(X^{-1})$ is positive on $(-1, 1)$. For $\epsilon < \alpha$, conditions **A1**, **A2** and **A3** result in

$$g(X) \ge C(X) - (\epsilon + 2\alpha X^{n-m}) \sum_{i=0}^{m-1} |X|^i + X^m n^{-1/4} \sigma_{n-m+1}(X) , \tag{8.12}$$

for all $|X| \le 1$. Since $(1-X^s)C(X) = (1-X^m)Q(X)$, we see that for n large enough and all $|X| \le 2^{-1/m}$,

$$(1-X^s)g(X) \ge (1-X^m)Q(X) - (\epsilon + 4\alpha 2^{-n/m}) \sum_{i=0}^{s-1} |X|^i \ge \frac{1}{2}Q(X) - 2s\epsilon$$

is positive for $\epsilon < Q(1)/(8s)$. Since $C(X) \ge \frac{m}{2s}Q(X)$ when $|X| \in [2^{-1/m}, \delta^{1/n}]$, it follows from (8.12) that

$$g(X) \ge m\left[\frac{Q(X)}{2s} - \epsilon - 2\alpha\delta^{1-m/n}\right]$$

is positive for any such X, provided $\epsilon < Q(1)/(8s)$, $2\alpha\delta^{1/2} < Q(1)/(8s)$ and n is large enough. Finally, for large n, if $|X| \in [\delta^{1/n}, 1]$ then $X^m n^{-1/4}\sigma_{n-m+1}(X) \ge n^{1/8}$ (see (8.8)). Since $m = o(\log n)$, the positivity of $g(X)$ for such X is a direct consequence of (8.12). $\square$

8.2 Proof for k and n odd

In this section we sketch the modifications to the argument of the previous section that are required for the case where k and n are odd. We will specify an event occurring with probability at least $n^{-b-o(1)}$ that forces $k-1$ simple zeros in $(0,1)$, one simple zero in $(-\infty,-1)$, and no other real zeros. Fix positive δ, integer r and $\epsilon > 0$, and define r_i, $\rho = \rho_n$ and $m = m_n$ as in Section 8.1. Define the polynomials

$$\begin{aligned} B(x) = \sum_{i=0}^{m-1} b_i x^i \quad &:= \quad (1 + x^s + x^{2s} + \ldots + x^{r_1 - s})Q(x) + (x^{r_1} + x^{r_1+s} + \ldots + x^{r_2-s})R(x) \\ &\qquad + (x^{r_2} + x^{r_2+s} + \ldots + x^{r_3-s})Q(x) + (x^{r_3} + x^{r_3+s} + \ldots + x^{r_4-s})R(x) \\ &\qquad + \ldots + (x^{r_{k-1}} + x^{r_{k-1}+s} + \ldots + x^{m-s})Q(x), \\ C(X) = \sum_{i=0}^{m} c_i X^i \quad &:= \quad (1 + X^s + X^{2s} + \ldots + X^{r_1-s})Q(X) \\ &\qquad + \alpha X^{r_1} + X(X^{r_1} + X^{r_1+s} + \ldots + X^{m-s})Q(X) \end{aligned}$$

the coefficients of which are in $\{\alpha, \beta\}$. Let $\mathcal{B}_n$ denote the event that the following hold:

B1 $|a_i - b_i| < \epsilon$ for $i = 0, 1, \ldots, m-1$

B2 $|a_{n-i} - c_i| < \epsilon$ for $i = 0, 1, \ldots, m$

B3 $a_m + a_{m+1}x + \ldots + a_{n-m-1}x^{n-2m-1} > n^{-1/4}\sigma_{n-2m}(x)$ for all $x \in \mathbb{R}$

B4 $|a_i| < n^\rho$ for $i = 0, 1, \ldots, n$.

Note that the degree of $C(X)$ is one larger than in Section 8.1. This ensures that the "middle polynomial" in condition **B3** has even degree, so that Theorem 1.3 applies to it. Hence, similarly to the proof of Lemma 8.1, one has that the event $\mathcal{B}_n$ occurs for odd n with probability exceeding $n^{-b-o(1)}$.

For all small enough $\delta > 0$, large enough r and small enough ϵ, the argument of the proof of Lemma 8.2, using the shape of $B(x)$, shows that if the coefficients of $f(x)$ are in $\mathcal{B}_n$ then $f(x)$ has exactly $k-1$ zeros in $[0,1]$, all simple, and no zeros in $[-1,0]$. We next prove that the function $F(X) := (1 - X^s)X^n f(1/X)$ satisfies

- $F(X) > 0$ for $X \in (0,1)$
- $F(X) > 0$ for $X \in [-\delta^{1/r_1}, 0)$
- $F'(X) > 0$ for $X \in [-(1-\delta)^{1/r_1}, -\delta^{1/r_1}]$
- $F(X) < 0$ for $X \in [-2^{-1/m}, -(1-\delta)^{1/r_1}]$
- $F(X) < 0$ for $X \in (-1, -2^{-1/m}]$

These will imply that $f(x)$ has a simple zero in $(-\infty,-1)$, and no other zeros with $|x| > 1$. Together with the $k-1$ simple zeros in $[0,1]$, this will bring the total number of zeros to k.

First, a proof analogous to that of (8.4) shows that there exists $c(r,\delta)$ finite, such that for $|X| \le (1-\delta)^{1/r_1}$,

$$|F(X) - (1 - X^s)C(X)| \le c(r,\delta)\epsilon. \tag{8.13}$$

The analogue of (8.5) is

$$(1-X^s)C(X) = (1-X^{r_1}+X^{r_1+1}-X^{m+1})Q(X) + \alpha X^{r_1}(1-X^s). \tag{8.14}$$

Suppose $X \in [0,(1-\delta)^{1/r_1}]$. Then (8.14) implies

$$(1-X^s)C(X) \geq (1-X^{r_1})Q(X) \geq \delta Q(X) > c(r,\delta)\epsilon$$

if ϵ is small enough, so $F(X)>0$ by (8.13).

Suppose $X \in [-\delta^{1/r_1},0]$. Then (8.14) implies

$$(1-X^s)C(X) \geq (1-3\delta)Q(X)$$

so $F(X)>0$ by (8.13) assuming suitable δ and ϵ.

Suppose $X \in [-(1-\delta)^{1/r_1},-\delta^{1/r_1}]$. The analogues of (8.9) and (8.10) are

$$\left|F'(X) - \frac{d}{dX}\left[(1-X^s)C(X)\right]\right| \leq c(\delta,r)\epsilon$$

and, with r_1, s even,

$$\begin{aligned}\frac{d}{dX}\left[(1-X^s)C(X)\right] &= \left(1-X^{r_1}+X^{r_1+1}\right)Q'(X) + \left[-r_1X^{r_1-1}+(r_1+1)X^{r_1}\right]Q(X)\\ &\quad + \alpha r_1 X^{r_1-1}(1-X^s) - \alpha s X^{r_1+s-1} - o(1)\\ &\geq -3M' + r_1\delta Q(X) - \alpha r_1(1-\delta^{s/r_1}) - o(1)\\ &\geq (r_1\delta/2)Q(X),\end{aligned}$$

in which the last inequality holds for r sufficiently large. Hence for ϵ small enough, $F'(X)$ will be positive.

Suppose $X \in [-2^{-1/m},-(1-\delta)^{1/r_1}]$. Then, for r sufficiently large, $1-X^{r_1}+X^{r_1+1}-X^{m+1} \leq -(1/2-3\delta)/2$ and $\alpha X^{r_1}(1-X^s) = O(s/r_1)$. For δ small and r large (8.14) thus implies that

$$(1-X^s)C(X) \leq -(1/2-3\delta)Q(X)/2 + O(s/r_1) \leq -Q(X)/8$$

Although (8.13) is no longer valid, we may apply **B3** to deduce

$$\begin{aligned}F(X)-(1-X^s)C(X) &\leq \epsilon s + (1-X^s)(a_mX^{n-m}+a_{m+1}X^{n-m-1}+\ldots+a_{n-m-1}X^{m+1})\\ &\quad + m(\alpha+\epsilon)X^{n-m+1}\\ &\leq \epsilon s + m(\alpha+\epsilon)2^{-(n-m+1)/m} \leq 2\epsilon s\end{aligned}$$

since $m2^{-n/m}\to 0$. Hence $F(X)<0$ if we take first δ small then r large and finally ϵ small.

Similarly, for r large enough, if $X \in [(1-\delta)^{1/r_1},2^{-1/m}]$, then

$$(1-X^s)C(X) \geq (1-X^m)XQ(X) \geq Q(X)/3,$$

and

$$F(X)-(1-X^s)C(X) \geq -\epsilon s - m(\alpha+\epsilon)X^{n-m+1} \geq -2\epsilon s$$

implying that $F(X)>0$ in this interval.

Both the proof that $F(X)>0$ on $[2^{-1/m},1)$ and the proof that $F(X)<0$ on $(-1,-2^{-1/m}]$ parallel the proof in Section 8.1 that $f(x)>0$ for $x\in[2^{-1/m},1)$. □

9 Proof of Proposition 1.5

In view of the upper bound $q_{n,k} \leq n^{-b+o(1)}$ of Section 7, it suffices to provide a lower bound on the probability of the event considered in Proposition 1.5. To this end, partitioning and shrinking the U_i if necessary, we may assume that $m_1 = \ldots = m_k = 1$, and that the closures of the U_i avoid both 1 and -1. Let $\delta \in (0, 1/3)$ then be such that each of the U_i is contained either in $(-1+\delta, 1-\delta)$ or its image under the map $\mathrm{inv}(x) = x^{-1}$. Let r be the number of U_i of the former type and $s = k - r$ the number of those of the latter type. Let $\mathbf{S} = (-\eta/2, \eta/2)$ for $\eta > 0$ as in the statement of the proposition. Fix the polynomials $B(x) = \sum_{i=0}^{r} b_i x^i \in \mathbf{S}[x]$ and $C(X) = \sum_{i=0}^{s} c_i X^i \in \mathbf{S}[X]$ with coefficients in $\mathbf{S}$, such that $B(x)$ has r real zeros, one in each of the U_i that are contained in $(-1, 1)$ whereas $C(X)$ has s real zeros, one in $\mathrm{inv}(U_i)$ for each U_i contained in $(-\infty, -1) \cup (1, \infty)$. Without loss of generality we can set $b_r > 0$ and $c_s > 0$. Let $\rho_n = 5/p_n$ for $p_n \uparrow \infty$ such that $E|a_i|^{p_n} \leq n$ (these differ from the quantities defined in Section 5). Define the even integer $m = m_n = 2\lfloor \rho_n \log n / |\log(1-\delta)| \rfloor$ depending on n. For fixed $\epsilon \in (0, \eta/11)$, consider the event $\mathcal{C}_n$ that all of the following are satisfied:

C1 $\quad |a_i - b_i| < \epsilon$ for $0 \leq i \leq r$, $\quad |a_{r+i} - 9\epsilon 1_{i \text{ even}}| < \epsilon$ for $0 < i < m$.

C2 $\quad |a_{n-i} - c_i| < \epsilon$ for $0 \leq i \leq s$, $\quad |a_{n-s-i} - 9\epsilon 1_{i \text{ even}}| < \epsilon$ for $0 < i < m$.

C3 $\quad a_{m+r} + a_{m+r+1}x + \ldots + a_{n-s-m}x^{n-k-2m} > 0$ for all $x \in \mathbb{R}$.

C4 $\quad |a_i| < n^{\rho_n}$ for $0 \leq i \leq n$.

Proposition 1.5 is an immediate consequence of the following two lemmas.

Lemma 9.1 *For any fixed $B(x)$, $C(X)$ with coefficients in $\mathbf{S}$ and positive $\epsilon < \eta/11$, the probability of the event $\mathcal{C}_n$ is at least $n^{-b-o(1)}$.*

Lemma 9.2 *For fixed $B(x)$ and $C(X)$, if $\epsilon > 0$ is sufficiently small and n sufficiently large, then any polynomial $f(x) = \sum_{i=0}^{n} a_i x^i$ satisfying the conditions of $\mathcal{C}_n$ has exactly k real zeros, one in each of the U_i intervals.*

Proof of Lemma 9.1: Note that $P(a \in G) > 0$ for any open subset G of $(-\eta, \eta)$ (by our assumption about the support of the law of a_i). Hence each coefficient condition in **C1** and **C2** is satisfied with probability at least c for some $c > 0$ depending only on $B(x)$, $C(X)$ and ϵ. We continue along the lines of the proof of Lemma 8.1 (taking now $n' = n - k - 2m + 1$ and $\gamma_{n'} = 0$). □

Proof of Lemma 9.2: Our choice of $\rho = \rho_n$ and $m = m_n$ guarantees that for any $l < \infty$, $m^l n^\rho (1-\delta)^m \to_{n\to\infty} 0$. Consequently, by **C1** and **C4**, for some $\kappa_0 = \kappa_0(\delta)$, all $\epsilon > 0$, $n > n_0$ for some $n_0 = n_0(\delta, \varepsilon)$ large enough and $|x| \leq (1-\delta)$

$$\begin{aligned} |f(x) - B(x)| &\leq 10\epsilon(1 + |x| + \ldots + |x|^{m+r-1}) + n^\rho(|x|^{m+r} + |x|^{m+r+1} + \ldots) \\ &\leq (10\epsilon + n^\rho(1-\delta)^{m+r})/\delta \leq \kappa_0 \epsilon \,. \end{aligned}$$

Hence if ϵ is small enough and n large enough, f must have at least as many zeros as $B(x)$ within $(-1+\delta, 1-\delta)$. On the other hand, $B^{(r)}(x)$ is a positive constant, and for $x \in (-1+\delta, 1-\delta)$,

$$\begin{aligned} |f^{(r)}(x) - B^{(r)}(x)| &\leq 10\epsilon \sum_{i=r}^{m+r-1} i^r |x|^{i-r} + n^\rho \sum_{i=m+r}^{\infty} i^r |x|^{i-r} \\ &\leq 10\epsilon \sum_{i=r}^{\infty} i^r (1-\delta)^{(i-r)} + n^\rho \sum_{i=m+r}^{\infty} i^r (1-\delta)^{(i-r)}, \end{aligned}$$

which again can be made arbitrarily small by shrinking ϵ. So, we can and shall assume $f^{(r)}(x) > 0$ in $(-1+\delta, 1-\delta)$. By Rolle's Theorem, this bounds the number of real zeros in $(-1+\delta, 1-\delta)$ by r, so $f(x)$ has exactly r zeros in $(-1+\delta, 1-\delta)$. Moreover, taking $\epsilon > 0$ such that $|B(x)| > \kappa_0 \epsilon$ for all $|x| \le (1-\delta)$, $x \notin U_i$, $i = 1, \ldots, k$, implies that the constant sign of $f(x)$ between each adjacent pair of intervals U_i that are contained in $(-1+\delta, 1-\delta)$ is the same as the sign of $B(x)$ there. Hence $f(x)$ has exactly one zero in each of the r intervals U_i contained in $(-1,1)$. Similar arguments (using **C2** and **C4**) show that for some $\kappa_1 = \kappa_1(\delta)$ and all $|X| < (1-\delta)$,

$$|X^n f(X^{-1}) - C(X)| \le \kappa_1 \epsilon,$$

with the s-th derivative of the polynomial $X^n f(X^{-1})$ made positive throughout $|X| < (1-\delta)$ by shrinking ϵ. Recall that $C(X)$ has exactly one zero in each of the intervals $\mathrm{inv}(U_i)$ for the U_i contained in $(-\infty, -1) \cup (1, \infty)$. Thus, for small enough ϵ, the same property holds for the s zeros of $X^n f(X^{-1})$ within $|X| < (1-\delta)$.

It thus remains to show that $x^r f(x) > 0$ for $(1-\delta) \le |x| \le (1-\delta)^{-1}$. Since $2r + m$ is an even integer, we have by condition **C3** that for all $x \in \mathbb{R}$,

$$x^r \left(\sum_{i=m+r}^{n-s-m} a_i x^i \right) = x^{2r+m} \left(\sum_{i=0}^{n-k-2m} a_{m+r+i} x^i \right) \ge 0\,. \tag{9.1}$$

It is easy to check that

$$h_m(x) := 8x^2 \sum_{j=0}^{(m-4)/2} x^{2j} - |x| \sum_{j=0}^{(m-2)/2} x^{2j} \ge 0\,,$$

for all even $m \ge 4$ and $2/3 \le |x| \le 3/2$. Consequently, for $(1-\delta) \le |x| \le (1-\delta)^{-1}$ and $\delta < 1/3$, by condition **C1**,

$$x^r \left(\sum_{i=r+1}^{r+m-1} a_i x^i \right) = x^{2r} \left(\sum_{i=1}^{m-1} a_{r+i} x^i \right) \ge \epsilon x^{2r} h_m(x) \ge 0,, \tag{9.2}$$

whereas for $r + (n-s) = 2r + n - k$ an even integer, by condition **C2**,

$$x^r \left(\sum_{i=n-s-m+1}^{n-s-1} a_i x^i \right) = x^{r+(n-s)} \left(\sum_{i=1}^{m-1} a_{n-s-i} x^{-i} \right) \ge \epsilon x^{2r+n-k} h_m(x^{-1}) \ge 0. \tag{9.3}$$

Next note that for sufficiently small $\epsilon > 0$, the polynomial $\sum_{i=0}^{r} a_i x^i$ has a positive leading coefficient and no zeros for $|x| \ge (1-\delta)$, so $x^r (\sum_{i=0}^{r} a_i x^i) > 0$ for all $|x| \ge (1-\delta)$. Similarly, $x^{-(n-s)} \sum_{i=n-s}^{n} a_i x^i$ is then a polynomial with positive constant coefficient and no zeros for $|x| \le (1-\delta)^{-1}$. With $r + (n-s)$ an even integer, it follows that $x^r (\sum_{i=n-s}^{n} a_i x^i) \ge 0$ for $|x| \le (1-\delta)^{-1}$. In view of (9.1)–(9.3), we find that $x^r f(x) > 0$ for $(1-\delta) \le |x| \le (1-\delta)^{-1}$. $\square$

References

[Ad] Adler, R. J. *An introduction to continuity, extrema, and related topics for general Gaussian processes.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, **12**. Institute of Mathematical Statistics, Hayward, CA, 1990.

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

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

[CS] Csörgő, M. and Révész, P. *Strong aprroximations in probability and statistics.* Academic Press, New York, 1981.

[DZ] Dembo, A; Zeitouni, O. *Large Deviations Techniques and Applications, second ed.* Springer-Verlag, New York, 1998.

[EK] Edelman, A.; Kostlan, E. *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.

[EO] Erdős, Paul; Offord, A. C. *On the number of real roots of a random algebraic equation.* Proc. London Math. Soc. **6** (1956), 139–160.

[Fa] Farahmand, K. *Topics in random polynomials.* Pitman research notes in mathematics series **393**. Longman, harlow, 1998.

[GR] Gradshteyn, I. S.; Ryzhuk, I. M. *Tables of integrals, series, and products*, 4-th ed. Academic Press, San Diego, 1980.

[IM1] Ibragimov, I. A.; Maslova, N. B. *The average number of zeros of random polynomials.* Vestnik Leningrad. Univ. **23** (1968), 171–172.

[IM2] Ibragimov, I. A.; Maslova, N. B. *The mean number of real zeros of random polynomials. I. Coefficients with zero mean.* Theor. Probability Appl. **16** (1971), 228–248.

[IM3] Ibragimov, I. A.; Maslova, N. B. *The mean number of real zeros of random polynomials. II. Coefficients with a nonzero mean.* Theor. Probability Appl. **16** (1971), 485–493.

[IM4] Ibragimov, I. A.; Maslova, N. B. *The average number of real roots of random polynomials.* Soviet Math. Dokl. **12** (1971), 1004–1008.

[IZ] Ibragimov, I. A.; Zeitouni, O. *On roots of random polynomials.* Trans. American Math. Soc. **349** (1997), 2427–2441.

[Ja] Jamrom, B. R. *The average number of real zeros of random polynomials.* Soviet Math. Dokl. **13** (1972), 1381–1383.

[Ka1] Kac, M. *On the average number of real roots of a random algebraic equation.* Bull. Amer. Math. Soc. **49**, (1943), 314–320. Erratum: Bull. Amer. Math. Soc. **49**, (1943), 938.

[Ka2] Kac, M. *On the average number of real roots of a random algebraic equation. II.* Proc. London Math. Soc. **50**, (1949), 390–408.

[KMT] Komlós, J., Major, P. and Tusnády, G. *An approximation of partial sums of independent R.V.'s and the sample D.F. II.* Z. Wahr. verw. Gebiete **34** (1976), 35–58.

[LO1] Littlewood, J. E.; Offord, A. C. *On the number of real roots of a random algebraic equation. I.* J. London Math. Soc. **13**, (1938), 288–295.

[LO2] Littlewood, J. E.; Offord, A. C. *On the number of real roots of a random algebraic equation. II.* Proc. Cambridge Philos. Soc. **35**, (1939), 133–148.

[LO3] Littlewood, J. E.; Offord, A. C. *On the number of real roots of a random algebraic equation. III.* Rec. Math. [Mat. Sbornik] N.S. **54**, (1943), 277–286.

[LS1] Logan, B. F.; Shepp, L. A. *Real zeros of random polynomials.* Proc. London Math. Soc. **18** (1968), 29–35.

[LS2] Logan, B. F.; Shepp, L. A. *Real zeros of random polynomials. II.* Proc. London Math. Soc. **18** (1968), 308–314.

[Ma1] Maslova, N. B. *The variance of the number of real roots of random polynomials.* Teor. Verojatnost. i Primenen. **19** (1974), 36–51.

[Ma2] Maslova, N. B. *The distribution of the number of real roots of random polynomials.* Theor. Probability Appl. **19** (1974), 461–473 (1975).

[PS] Poonen, B.; Stoll, M. *The Cassels-Tate pairing on polarized abelian varieties.* Annals of Math. **150** (1999), 1109–1149.

[Sh] Shao, Q.M. *A Gaussian correlation inequality and its applications to the existence of small ball constant.* Preprint, (1999).

[St] Stevens, D. C. *The average number of real zeros of a random polynomial.* Comm. Pure Appl. Math. **22** (1969), 457–477.

[Sto] Stout, W. F. *Almost sure convergence.* Academic Press, New York, 1974.

[Str] Strassen, V. *An invariance principle for the law of the iterated logarithm.* Z. Wahrscheinlichkeitstheorie verx. Gebiete **3**, (1964), 211-226.

[To] Todhunter, I. *A history of the mathematical theory of probability.* Stechert, New York, 1931.

[Wa] Wang, You Jing. *Bounds on the average number of real roots of a random algebraic equation.* Chinese Ann. Math. Ser. A **4** (1983), 601–605. An English summary appears in Chinese Ann. Math. Ser. B **4** (1983), 527.

[Wi] Wilkins, J. Ernest, Jr. *An asymptotic expansion for the expected number of real zeros of a random polynomial.* Proc. Amer. Math. Soc. **103** (1988), 1249–1258.