J. Eur. Math. Soc. 28, 2261–2305 (2026) DOI 10.4171/JEMS/1530

Marcus Michelen · Julian Sahasrabudhe

Central limit theorems and the geometry of polynomials

Received 31 July 2022; revised 5 June 2024

Abstract. Let $X \in \{0, \dots, n\}$ be a random variable with mean μ, standard deviation σ, and let $f_X(z) = \sum_k \mathbb{P}(X = k) z^k$ be its probability generating function. Pemantle conjectured that if σ is large and f_X has no roots close to $1 \in \mathbb{C}$, then X must be approximately normal. We completely resolve this conjecture in the following strong quantitative form, obtaining sharp bounds. If $\delta = \min_\zeta |\zeta - 1|$ over the complex roots ζ of f_X, and $X^* := (X - \mu)/\sigma$, then $\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O((\log n)/(\delta\sigma))$, where $Z \sim \mathcal{N}(0, 1)$ is a standard normal variable. This gives the best possible version of a result of Lebowitz, Pittel, Ruelle and Speer. We also show that if f_X has no roots with small *argument*, then X must be approximately normal, again in a sharp quantitative form: if we set $\delta = \min_\zeta |\arg(\zeta)|$, then $\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O(1/(\delta\sigma))$. Using this result, we answer a question of Ghosh, Liggett and Pemantle by proving a sharp multivariate central limit theorem for random variables with real-stable probability generating functions.

Keywords: geometry of polynomials, stable polynomials, strongly Rayleigh.

1. Introduction

In his influential paper on negative dependence, Pemantle [56] set out a list of desirable combinatorial properties for “the correct” definition of negatively dependent random variables and laid out a number of natural conjectures. In their celebrated paper, Borcea, Brändén and Liggett [14] provided such a definition by making a striking connection with the blossoming subject of real-stable polynomials; it turns out that the definition that Pemantle sought is best described in terms of the zeros of the associated probability generating function. For this, let $X \in \{0, \dots, n\}^d$ be a random variable, let

$$f_X(z_1, \dots, z_d) := \sum_{(i_1, \dots, i_d)} \mathbb{P}(X = (i_1, \dots, i_d)) z_1^{i_1} \cdots z_d^{i_d},$$

Marcus Michelen: Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 851 South Morgan Street, Chicago, IL 60607-7045, USA; michelen@uic.edu, michelen.math@gmail.com

Julian Sahasrabudhe: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK; jdrs2@cam.ac.uk

Mathematics Subject Classification 2020: 60F05 (primary); 30C10, 60E10 (secondary).

be its probability generating function and define

$$\mathbb{H} = \{(z_1, \dots, z_d) \in \mathbb{C}^d : \mathrm{Im}(z_i) > 0 \text{ for all } i\}$$

to be the *upper half-plane*.[1] A polynomial $f \in \mathbb{R}[z_1, \dots, z_d]$ is said to be *real-stable* if it has no roots ζ in the upper half-plane $\mathbb{H}$ and a random variable X is said to be *strong Rayleigh* if its probability generating function f_X is real-stable. Borcea, Brändén and Liggett showed that strong Rayleigh random variables admit a natural theory of negative dependence and provided many natural examples of strong Rayleigh distributions: spanning tree distributions, uniform random matching distributions in graphs and determinantal measures. In the years since this notion has been well studied and many further examples have been found [2, 40, 43, 44, 55, 57, 58, 74].

In addition to the connection with negative dependence, the theory of real-stable polynomials has had many recent successes, notably Borcea and Brändén's [12, 13] powerful classification of linear operators that preserve real-stability; its role in Marcus, Spielman, and Srivastava's spectacular proof of the Kadison–Singer conjecture [50]; and in Gurvits's surprising and simple proof of (and extensions of) the van der Waerden conjecture [31,32]; among others [3, 39, 49].

In this paper, one of our main motivations is to finish a program set in motion by Ghosh, Liggett and Pemantle [28] to show that if $X_n \in \{0, \dots, n\}^d$ is a sequence of random variables, with real-stable generating functions, then X_n tends to a multivariate Gaussian distribution, after centering and scaling, provided $\sigma_n \to \infty$. We will derive this theorem by first proving results on univariate polynomials, with much looser restrictions on the roots, and then "lifting" these results to the multivariate setting.

In the univariate setting, work on the connection between roots of polynomials and their coefficients reaches back (at least) to Cauchy's quantitative work on the fundamental theorem of algebra [16], but was perhaps first intensely studied by Littlewood and Offord [45–47], Szegő [70], Bloch and Póyla [11] and Schur [65] among others (see [51] for more discussion). To give a bit of flavor of these results, we mention only one such result from this vast literature that is most relevant for us here. In 1950, Erdős and Turán [24] proved that if $P(z) = \sum_{k=1}^n a_k z^k$ is a polynomial ($a_0 a_n \neq 0$) with sufficiently "flat" coefficients, meaning $(|a_0||a_n|)^{-1/2} \sum_{k=1}^n |a_k| = e^{o(n)}$, then the roots of P are approximately radially-equidistributed in the complex plane, meaning that each sector $\{z : \alpha \leqslant \arg(z) \leqslant \beta\}$, for $0 \leqslant \alpha < \beta \leqslant 2\pi$ contains roughly $n(\beta - \alpha)/2\pi$ roots. This result has been adapted to different settings [23] and generalized and sharpened several times [1, 10, 53]. For more details, we refer the reader to the lovely articles of Granville [30] and Soundararajan [69].

In this paper, we show that a substantial amount of information about the coefficients of a polynomial can be derived from its locus of zeros, if we additionally assume the polynomial is a probability generating function, which is to say, it has non-negative

[1] Throughout the paper, we will slightly abuse notation and write $X \in S$ for a random variable X and a set S as shorthand for "X takes values in the set S".

coefficients. A surprising first step in this direction is due to Lebowitz, Pittel, Ruelle and Speer [41], who showed that if, for each $n \geqslant 1$, $X_n \in \{0, \dots, n\}$ is a random variable for which f_{X_n} has no zeros in a neighborhood of $1 \in \mathbb{C}$, and $\sigma_n n^{-1/3} \to \infty$, then $(X_n - \mu_n)\sigma_n^{-1}$ tends weakly to a normal distribution (see also the 1979 work of Iagolnitzer and Souillard in the context of the Ising model [36]). Inspired by this advance, Pemantle ([51] and personal communication by R. Pemantle, 2017) was lead to conjecture that the variance condition in the theorem of Lebowitz, Pittel, Ruelle and Speer could be greatly improved.

Conjecture 1.1 (Pemantle, 2017). *For $\delta > 0$ and each $n \geqslant 1$, let $X_n \in \{0, \dots, n\}$ be a random variable with mean μ_n, standard deviation σ_n and for which the roots ζ of the probability generating function f_{X_n} satisfy $|\zeta - 1| \geqslant \delta$. Then $(X_n - \mu_n)\sigma_n^{-1} \to N(0, 1)$, provided $\sigma_n \to \infty$.*

In the recent work [51], the authors refuted this conjecture by showing that for any $C > 0$, there exist random variables $X_n \in \{0, \dots, n\}$ with $\sigma_n > C \log n$ that *are not* asymptotically normal and for which f_{X_n} has no roots in a neighborhood of $1 \in \mathbb{C}$. On the other hand, the authors also showed that Pemantle was right to suspect that the variance condition in the work of Lebowitz, Pittel, Ruelle and Speer could be significantly improved, by showing that it is sufficient to assume $\sigma_n > n^{\varepsilon}$, for any $\varepsilon > 0$.

Here, we completely resolve the conjecture of Pemantle by showing that a sufficient condition for convergence to a normal distribution is to have $\sigma_n(\log n)^{-1} \to \infty$. In fact, we prove a sharp quantitative version of this theorem that gives an optimal bound on the maximum discrepancy between a random variable X and a standard normal, based only on the distance of the closest root of f_X to $1 \in \mathbb{C}$.

Theorem 1.2. *Let $X \in \{0, \dots, n\}$ be a random variable with mean μ, standard deviation σ and probability generating function f_X and set $X^* = (X - \mu)\sigma^{-1}$. If $\delta \in (0, 1)$ is such that $|1 - \zeta| \geqslant \delta$ for all roots ζ of f_X, then*[2]

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O\Big(\frac{\log n}{\delta \sigma}\Big), \tag{1.1}$$

where $Z \sim N(0, 1)$.

We note that this immediately implies the following limit theorem for distributions with no roots close to $1 \in \mathbb{C}$.

Corollary 1.3. *For each $n \geqslant 1$, let $\delta_n \in (0, 1)$, and let $X_n \in \{0, \dots, n\}$ be a random variable with mean μ_n, standard deviation σ_n and probability generating function f_n. If $|\zeta - 1| \geqslant \delta_n$ for all roots ζ of f_n and*

$$\sigma_n \delta_n (\log n)^{-1} \to \infty,$$

then $(X_n - \mu_n)\sigma_n^{-1} \to N(0, 1)$ in distribution.

[2] The implicit constant can be taken to be 2^{3261}.

The condition on the standard deviation σ_n in Corollary 1.3 is sharp, both in terms of δ_n and in terms of n.

Our second result (and the main ingredient in the proof of the multivariate central limit theorem for strong Rayleigh distributions) says we can weaken the variance condition in Theorem 1.2 all the way to $\sigma_n \to \infty$ (the obvious[3] necessary condition) if we further assume that the sequence f_n has no roots in a small sector $\{z : |\arg(z)| < \delta\}$ containing the positive real axis. Again, we prove a sharp, quantitative version of this by obtaining an optimal bound on the discrepancy between a normal random variable $Z \sim N(0, 1)$ and a random variable X, based only on the smallest angle made by a root of f_X and the positive real axis.

Theorem 1.4. *Let $X \in \{0, \ldots, n\}$ be a random variable with mean μ, standard deviation σ and probability generating function f_X and set $X^* = (X - \mu)\sigma^{-1}$. If $\delta > 0$ is the minimum of $|\arg(\zeta)|$ over the roots ζ of f_X, then*[4]

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O\Big(\frac{1}{\delta\sigma}\Big),$$

where $Z \sim N(0, 1)$.

Theorem 1.4 immediately implies the following limit theorem for distributions where the smallest argument of a root *just* exceeds the reciprocal of the standard deviation.

Corollary 1.5. *For each $n \geqslant 1$, let $X_n \in \{0, \ldots, n\}$ be a random variable with mean μ_n, standard deviation σ_n and probability generating function f_n. If the roots ζ of f_n satisfy $|\arg(\zeta)| \geqslant \delta_n$ and*

$$\delta_n\sigma_n \to \infty,$$

then $(X_n - \mu_n)\sigma_n^{-1} \to N(0, 1)$, in distribution.

Again, as we shall see in Section 11, the condition on σ_n is sharp for all sequences $(\delta_n)_n$.

With Theorem 1.4 in hand, it is not hard to prove our multivariate central limit theorem for strong Rayleigh distributions, following a key observation of Ghosh, Liggett and Pemantle. To properly state this result, recall that if $X = (X_1, \ldots, X_d) \in \mathbb{R}^d$ is a random variable, its *covariance matrix* $A = A(X)$ is a symmetric, semi-definite matrix defined by

$$(A)_{i,j} := \mathbb{E} X_i X_j - \mathbb{E} X_i \, \mathbb{E} X_j,$$

and its *maximum variance* σ_n^2 is defined as the ℓ^2-operator norm of A. For a $d \times d$ positive semi-definite matrix A and $\mu \in \mathbb{R}^d$, we denote the Gaussian random variable with covariance matrix A and mean μ by $N(\mu, A)$.

[3]The law of $(X - \mu_n)\sigma_n^{-1}$ is supported on point masses of distance $\geqslant \sigma_n$ so we must have $\sigma_n \to \infty$ if the sequence approximates the continuous Gaussian distribution.

[4]The implicit constant can be taken to be 2^{3257}.

Motivated by a vast literature on central limit theorems and multivariate generating functions [5,6,15,27,59,60], Ghosh, Liggett and Pemantle proved a central limit theorem for strong Rayleigh distributions, in the case that the sequence maximum variances σ_n^2 grows sufficiently quickly, namely $\sigma_n n^{-1/3} \to \infty$. In particular, they proved that if $d \in \mathbb{N}$ and $X_n \in \{0,\dots,n\}^d$ is a sequence of strong Rayleigh distributions with covariance matrices $\{A_n\}$ that satisfy $\sigma_n^{-2} A_n \to A$ and $\sigma_n n^{-1/3} \to \infty$, then $(X_n - \mu_n)\sigma_n^{-1} \to N(0, A)$, weakly. They conclude their paper by asking for the best possible condition on the growth of σ_n and ask, in particular, if the condition $\sigma_n \to \infty$ is sufficient in their theorem. In [51], we made progress on this problem by showing that $\sigma_n > n^{\varepsilon}$ is sufficient for any $\varepsilon > 0$. Here, we are able to completely resolve the question of Ghosh, Liggett and Pemantle by showing that the obvious necessary condition $\sigma_n \to \infty$ is indeed sufficient.

Theorem 1.6. *For $d \in \mathbb{N}$ and each $n \geqslant 1$, let $X_n \in \{0,\dots,n\}^d$ be a random variable with covariance matrix A_n and maximum variance σ_n^2. If the probability generating functions of X_n are real-stable, $\sigma_n \to \infty$ and $\sigma_n^{-2} A_n \to A$, then*

$$(X_n - \mu_n)\sigma_n^{-1} \to N(0, A),$$

in distribution.

We can additionally prove a quantitative form of this theorem (in the spirit of Theorems 1.2 and 1.4), but we defer this more technical result to a later paper.

It is perhaps interesting to note that in this paper, we make essentially no use of the rich theory of stable polynomials and, as a result, our work here provides (what appears to be) a new and flexible tool-set for working with real-stable polynomials. To illustrate, in Section 10 we show that our method immediately implies a version of Theorem 1.6 for *Hurwitz stable polynomials* [35], a similar and well-studied notion [17,71,72] along with other polynomials satisfying a similar "half-plane property". Our methods are also of use beyond proving central limit theorems. In a subsequent paper [52], we use the tools from this paper to prove a close connection between the roots of f_X and the *variance* of X.

1.1. General forms of main theorems

While all of our theorems above have been stated for random variables taking values in $\{0,\dots,n\}$ (the form in which these conjectures were posed), it is not hard to see that our methods imply similar results for more general random variables if we move to a slightly more general framework.

Here we state a result from Section 12, our main technical theorem,[5] from which all others follow. Note the condition that X is integer-valued is not present at all.

[5] See Section 3.2 for the definition of cumulants and Section 12 for appropriate generalizations of f_X to non-integer-valued random variables.

Theorem 1.7. *For $\varepsilon \in (0,1)$ and $b \geqslant 0$, let X be a random variable with logarithmic potential $u_X(z) = \log|f_X(z)|$ and cumulant sequence $(\kappa_j)_j$. If u_X is harmonic in $B(1,\varepsilon)$, for all $0 \leqslant \theta_1 \leqslant \theta_2$ and $r > 0$ with $re^{i\theta_1}, re^{i\theta_2} \in B(1,\varepsilon)$, we have $u_X(re^{i\theta_1}) - u_X(re^{i\theta_2}) > -b$, and*

$$\sum_{j \geqslant 2} \frac{|\kappa_j|}{j!}\left(\frac{\varepsilon}{32}\right)^j > b,$$

then

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O\left(\frac{1}{\varepsilon\sigma}\right),$$

where $Z \sim N(0,1)$ is a standard normal variable.

To illustrate a more general application of Theorem 1.7, we highlight Theorem 12.2 here, which is a natural generalization of Theorem 1.4, from Section 12. We point out that there is an additional growth hypothesis on $f(z)$ appearing in this statement which was "invisible" before since it is automatically satisfied for polynomials.

Theorem 1.8. *For $\delta > 0$ and $\kappa > 0$, let $X \in \mathbb{R}$ be a random variable with mean μ, standard deviation σ and with probability generating function f_X. If f_X is defined on $\mathbb{R}_{\geqslant 0}$, is zero-free in $\{z : |\arg(z)| < \delta\}$ and satisfies*

$$|\log|f_X(z)|| = o(|z|^\kappa) \quad \text{and} \quad \left|\log\left|f_X\left(\frac{1}{z}\right)\right|\right| = o(|z|^\kappa),$$

as $z \to \infty$ with $|\arg(z)| < \delta$, then

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O\left(\frac{\max\{\delta^{-1},\kappa\}}{\sigma}\right),$$

where $Z \sim N(0,1)$ and $X^* := \sigma^{-1}(X-\mu)$.

Again, this theorem is sharp in the dependence on all of κ, σ and δ. It is not hard to extract a similar generalization of Theorem 1.2 in this setting.

1.2. Background

The use of roots to study combinatorial distributions has a long and distinguished history in mathematics and has provided many surprising connections, with the most classical instance coming from the connection between the location of the roots of the Riemann zeta function with the distribution of the primes.

In statistical physics, Lee and Yang [42, 73] drew a surprising and influential connection between the roots of polynomials and phase transitions in physical systems by showing that the zero-freeness of certain partition functions implies the non-existence of phase transitions.

In combinatorics, the roots of various polynomials associated with graphs and other combinatorial objects have been shown to have particular regions without zeros (see, for example, [26, 37, 61, 62]). The two most classical examples here are the striking theorem of Heilmann and Lieb [34], which says that the roots of the matching polynomial are *real*, and the classical theorem of Lee and Yang [42, 73] who showed that the roots of the probability generating function (i.e., the partition function), associated with the number of "up-spins" of the Ising model on a finite graph, always lie on the unit circle.

In some cases, zero-free regions have been used directly to prove central limit theorems. In the case of matchings, a line of results [29, 33, 64], starting with the work of Godsil [29] and culminating in the work of Kahn [38], used the Heilmann–Lieb theorem to give general results when the size of a random matching in a graph is approximately normal. In the case of the Ising model, the work of Lee and Yang was used to prove central limit theorems for the number of "up-spins" by Iagolnitzer and Souillard [36] and later by Lebowitz, Pittel, Ruelle and Speer [41]. In a similar vein, Scott and Sokal [66, 67], who built on the work of Shearer [68] and Dobrushin [22], showed a close connection between zero-free regions and the Lovasz local lemma, another core probabilistic tool.

The philosophy that appears to have emerged from these advances is that the roots of combinatorially defined objects often have special structure and admit particular zero-free regions. This observation was made explicit by Rota, who sought to give "combinatorial meaning" to the distributions of roots in these settings [63]. In this light, one could see our results as a general contribution to this program of Rota (his so-called *critical program*) by giving combinatorial meaning to the roots of a wide class of polynomials.

2. Outline of proof

Theorems 1.2 and 1.4 are proved in parallel and can be thought of as two consequences of the same general method. As such, in the discussion here, we are intentionally vague about which of these two theorems we are proving. Now, let $X \in \{0, \dots, n\}$ be a random variable with probability generating function f_X and consider the *characteristic function* of X, which is a relative of f_X and defined as $\psi_X(\xi) := f_X(e^{i\xi})$, where $\xi \in \mathbb{R}$. The relevant feature of the characteristic function is that it detects the closeness between two probability distributions: a sequence of random variables Y_n converge in distribution to the random variable Y if and only if the sequence of characteristic functions ψ_{Y_n} converge to the characteristic function of Y ψ_Y *point-wise*. Of course, our results here are quantitative, but this fact serves as some guide: to show that Y is approximately normal it is enough to show that $\psi_Y(\xi) \approx e^{-\xi^2/2}$, where $\psi_Z := e^{-\xi^2/2}$ is the characteristic function of the standard normal $Z \sim N(0, 1)$. With this in mind, it is natural to center and scale X, by writing $X^* := (X - \mu)\sigma^{-1}$ and then to consider the logarithm of ψ_{X^*}, due to the exponential form of ψ_Z. Indeed, we will be able to express $\log|\psi_{X^*}(\xi)|$ as

$$\log|\psi_{X^*}(\xi)| = \sum_{j \geqslant 2} a_j \sigma^{-j} \mathrm{Re}(i^j \xi^j),$$

where ξ is in a sufficiently small neighborhood of 0 and (a_j) is a sequence of real numbers.

It turns out that $a_2 = -\sigma^2/2$ and hence the first term of the series is $-\xi^2/2$, just as we saw in the exponent of ψ_Z. From this vantage, our task becomes clear: we need to show that $|a_j| \ll \sigma^j$ in order to have $\psi_{X^*} \approx e^{-\xi^2/2}$.

With our goal now laid out, we turn to consider the function $u(z) := \log|f_X(z)|$ in a region around $1 \in \mathbb{C}$ in the complex plane. Note that here we can quite naturally make use of our zero-free hypothesis: if f_X is zero-free in a region, then the function u is harmonic in this region. Now, while the fact that u is harmonic on a particular region is a useful property, it is far from enough to prove our main theorems;[6] we will additionally need to make particular use of the fact that f_X has positive coefficients, a property that we use in the form of "weak positivity" for the function u (see Section 3).

As we will see in Section 4, this notion of weak positivity interacts nicely with the harmonic property of u, to give us another "positivity" notion which we make heavy use of. For $b \geqslant 0$, $\varepsilon > 0$, we say that function u on $B(1, \varepsilon)$, with $u(z) = u(\overline{z})$, is b-*decreasing* if for all $0 < \theta_1 < \theta_2 < \delta$, we have

$$u(\rho e^{i\theta_1}) - u(\rho e^{i\theta_2}) \geqslant -b,$$

where the functions are defined. In Section 4, we prove Lemma 4.1, which is our main tool for showing that a function is b-decreasing. This lemma says that if u is a weakly positive, harmonic function on $S := \{z \in \mathbb{C} : R^{-1} < |z| < R, |\arg(z)| < \delta\}$, then the function u is b-decreasing if

$$\exp\Big(-\delta^{-1} \log\Big(\frac{R}{r}\Big)\Big) \max_{z} |u(z)| \leqslant \frac{b}{10}, \tag{2.1}$$

where the maximum is taken over the "ends" of S, defined as

$$S^* := \{z \in \mathbb{C} : |z| \in \{R^{-1}, R\}, \arg(z) \in [-\delta, \delta]\}.$$

Without going into details, one can already see two important features of (2.1). Firstly, if u is harmonic and weakly decreasing in an *entire* sector $\{z : \arg(z) \in [-\delta, \delta]\}$, then the left-hand side of (2.1) can be taken to be arbitrarily small (by letting $R \to \infty$), and so we learn that u is 0-decreasing, which perhaps should strike the reader as a reasonably strong property. Secondly, we note the exponential dependence on the width δ of the sector. This ultimately accounts for the factor of $\log n$ that appears in Theorem 1.2.

With this tool in place, we turn to show how to use the b-decreasing hypothesis to get some control of the sequence $(a_j)_{\geqslant 2}$. In a series of steps, we work towards Lemma 8.1 and its important corollary, Corollary 9.3. These results are perhaps the main technical contributions of this paper, and their proof consumes Sections 5–8. Lemma 8.1 says that if u is a b-decreasing, weakly positive, harmonic function on S, then the associated characteristic function ψ of X^* must look like

$$\psi(\xi) = \exp\Big(-\frac{\xi^2}{2} + R(\xi)\Big), \tag{2.2}$$

[6]In fact, as we discussed in [51], the results of Lebowitz, Pittel, Ruelle and Speer are actually *sharp* if one generalizes their theorem to polynomials that have negative coefficients. Thus, we *must* use the non-negativity hypothesis in an essential way.

where the "remainder term" $R(\xi)$ is controlled by

$$|R(\xi)| \leqslant c\frac{|\xi|^3}{\delta\sigma}$$

for all $\xi \in \mathbb{R}$ satisfying $|\xi| < c_2\delta\sigma$. Of course, the reader should interpret (2.2) as saying that "ψ looks like the characteristic function of a standard normal, up to the remainder-term R".

The proof of (2.2) is carried out in three main steps. The first step is proved in Section 5 where we prove an important supporting result, Lemma 5.1, that allows us to compare the maximum of $u_0 := u - \mu \log|z|$, a re-normalized form of u, to a particular function $\varphi_{\gamma,b}(z)$ (defined in Section 3) which is both harmonic and *positive* on a region containing $1 \in \mathbb{C}$.

In Section 6, we use this "comparison" lemma to prove Lemma 6.1, which tells us that the sequence $(a_j)_{j\geqslant 2}$ has nice decay properties; for every $L \geqslant 2$, we have that

$$\frac{\sum_{j\geqslant L}|a_j|\varepsilon^j}{\sum_{j\geqslant 2}|a_j|\varepsilon^i} \leqslant C\cdot 2^{-L}, \tag{2.3}$$

provided

$$\sum_{j\geqslant 2}|a_j|\varepsilon^j > b,$$

where C is a large, but absolute, constant and $\varepsilon \approx \delta$.

We stress that (2.3) is a major step towards proving Lemma 8.1 and indeed Sections 5 and 6 are probably the most pivotal in the paper. However, (2.3) is not quite enough. Roughly speaking, (2.3) says that we have quite a bit of the "mass" of the sequence $(a_j)_{j\geqslant 2}$ is focused on the early terms of this sequence. We actually need to show that "most" of the mass is on the *second term*, $a_2 = -\sigma^2/2$.

For this next step, carried out in Section 7, we prove Lemma 7.5, which says that if u is weakly positive and harmonic around $1 \in \mathbb{C}$, *and* $|a_j|$ is large for some small $j \geqslant 2$, then $|a_2|$ must also be large. This allows us to control the magnitude of each of the terms $|a_j|$ relative to the value of $|a_2|$. Applying Lemmas 6.1 and 7.5 in sequence allows us to deduce (2.2).

Now, while (2.2) tells us that the characteristic function ψ of X^* is roughly like the characteristic function of a standard normal, we really care about showing that the *distribution* of X^* is close to the distribution of a standard normal. For this, we need an appropriate "Fourier inversion" step. This step is carried out in Section 9, just before we go on to deduce Theorems 1.2 and 1.4.

In Section 10, we turn to use the results developed in previous chapters to prove our multivariate central limit theorem for strong Rayleigh distributions. This is achieved by using first a fundamental observation of Ghosh, Liggett and Pemantle that says that if $X \in \{0,\ldots,n\}^d$ is a random variable with real-stable generating function, then the characteristic functions of the one-dimensional projections $\langle X, v\rangle$, where $v \in \mathbb{Z}^d_{\geqslant 0}$, have no roots in a small sector. Theorem 1.4 then allows us to show all of these projections are approximately normal. We then use a strong version of the Cramér–Wold theorem to lift this information to deduce that X itself must be approximately normal.

In Section 11, we give examples, demonstrating the tightness of our results. Finally, in Section 12 we briefly discuss how the main results of this paper can be generalized to go beyond polynomials to prove sharp results for power series and more general analytic functions.

3. Definitions and basic properties

In this section, we fix a few notations and introduce the central objects of our proof.

Throughout, we use the notations $\mathbb{R}_{\geqslant 0}$, $\mathbb{R}_{\leqslant 0}$ and so on to denote the non-negative reals and non-positive reals, respectively, and extend these definitions in the obvious way to $\mathbb{Z}$. If $z \in \mathbb{C}$, we write $z = re^{i\theta}$, where $r > 0$ and $\theta \in [-\pi, \pi]$, and then define the *argument of* z to be $\arg(z) = \theta$. For $-\pi \leqslant \beta < \alpha \leqslant \pi$, we define the *sector*

$$S(\alpha, \beta) := \{z \in \mathbb{C} \setminus \{0\} : \alpha \leqslant \arg z \leqslant \beta\}$$

and $S(\alpha) := S(-\alpha, \alpha)$. For $R \geqslant 1$ and $\varepsilon > 0$, we define the *truncated sector*

$$S_R(\varepsilon) := \{z \in \mathbb{C} : |z| \in [R^{-1}, R] \text{ and } \arg(z) \in [-\varepsilon, \varepsilon]\}$$

and define

$$S_R^*(\varepsilon) := \{z \in \mathbb{C} : |z| \in \{R^{-1}, R\} \text{ and } \arg(z) \in [-\varepsilon, \varepsilon]\}$$

to be the *ends* of the sector $S_R(\varepsilon)$. We also use the notation $S_R(\alpha, \beta) = S(\alpha, \beta) \cap S_R(\pi)$ in a similar way and use the (standard) notation $\partial\Omega$ to denote the boundary of a region $\Omega \subseteq \mathbb{C}$.

3.1. The logarithmic potential

If $X \in \{0, \dots, n\}$ is a random variable,[7] define

$$f_X(z) = \mathbb{E}z^X = \sum_{k=0}^{n} \mathbb{P}(X = k)z^k$$

to be its probability generating function, $\mu = \mathbb{E}X$ for its mean and $\sigma^2 = \mathrm{Var}[X]$ for its variance. Also note that $f_X(1) = 1$. Now define

$$u(z) = u_X(z) := \log|f_X(z)|$$

[7] For more general random variables, we will assume that X has an exponential moment and we define f_X for z in a neighborhood of 1 by choosing a branch of the logarithm (see Section 12 for more details).

to be the *logarithmic potential* of X and observe that if $f_X(z)$ is zero-free in an open set $\Omega \subseteq \mathbb{C}$, then u is *harmonic* on Ω. This key connection will allow us to exploit the theory of harmonic functions in certain regions of $\mathbb{C}$. We will say that a function u on Ω is *symmetric on* Ω if

$$u(z) = u(\overline{z})$$

for all z with $z, \overline{z} \in \Omega$. Of course, the logarithmic potential $u_X(z)$ is symmetric as f_X is a polynomial with real coefficients and so

$$u(z) = \log|f_X(z)| = \log|\overline{f_X(z)}| = \log|f_X(\overline{z})| = u(\overline{z}).$$

A third key property is particular to the fact that f_X is a probability generating function; that is, it is a polynomial with non-negative coefficients. We say that a function u is *weakly positive* on Ω if

$$u(|z|) - u(z) \geqslant 0$$

for all $z \neq 0$ with $z, |z| \in \Omega$. Weak positivity of u_X follows by taking the logarithm of both sides of the inequality

$$|f_X(z)| = |\mathbb{E}z^X| \leqslant \mathbb{E}|z|^X = |f_X(|z|)|.$$

We also note a useful expression of u_X in terms of the roots $\{\zeta\}$ of f_X

$$u_X(z) = \sum_{|\zeta|<1} \log\left|1 - \frac{\zeta}{z}\right| + \sum_{|\zeta|\geqslant 1} \log\left|1 - \frac{z}{\zeta}\right| + c_X + N_X \log|z|,$$

where c_X is defined so that $u_X(1) = \log|f_X(1)| = 0$ and N_X is the number of roots of f_X with $|\zeta| < 1$.

3.2. The exponential scale

We shall often work with the function $u = u_X$ on an "exponential scale" by defining $U(w) := u(e^w)$. Note that $U(w)$ is harmonic when u is (in the appropriate domains) and is also symmetric since

$$U(\overline{w}) = u(e^{\overline{w}}) = u(\overline{e^w}) = u(e^w) = U(w).$$

The importance of this form is made clear by Lemma 3.1; the Taylor expansion of $U(w)$ at $w = 0$ reveals the *cumulants* of X, which we denote by $(\kappa_j)_{j\geqslant 1}$. We do not need to draw on much external information here about this important sequence, but we do need to note that the first and second cumulants are familiar probabilistic quantities. Indeed,

$$\kappa_1 = \frac{\mathrm{d}}{\mathrm{d}w} u(e^w)\Big|_{w=0} = \mu, \tag{3.1}$$

and

$$\kappa_2 = \frac{\mathrm{d}^2}{\mathrm{d}w^2} u(e^w)\Big|_{w=0} = \sigma^2, \tag{3.2}$$

which exist under the condition that u is harmonic in a neighborhood of $1 \in \mathbb{C}$.

Since our interest is in central limit theorems, when working with u it is often useful to "subtract out" the term corresponding to μ. In particular, define

$$u_0(z) := u(z) - \mu \log|z|, \tag{3.3}$$

and correspondingly define $U_0(w) := u_0(e^w)$. If u is an arbitrary function (that is, not necessarily coming from a random variable), we may define u_0 and U_0 in the same way, by simply taking

$$\mu := \frac{\mathrm{d}}{\mathrm{d}w} u(e^w)\Big|_{w=0},$$

in the case that this quantity exists.

Lemma 3.1. *For $\varepsilon \in (0, 1/2)$, let u be a symmetric harmonic function on $B(1, 2\varepsilon) \subseteq \mathbb{C}$ with $u(1) = 0$. Then if U, U_0 are defined as above, we may express*

$$U(w) = \sum_{j \geqslant 1} a_j \mathrm{Re}(w^j) \tag{3.4}$$

and

$$U_0(w) = \sum_{j \geqslant 2} a_j \mathrm{Re}(w^j) \tag{3.5}$$

for all $w \in B(0, \varepsilon)$. Here the a_j are real numbers and $j!a_j = \kappa_j$, where κ_j is the j-th cumulant.

Proof. First note that $U(0) = u(1) = 0$ and that $U(w)$ is harmonic and symmetric in a ball $B(0, \varepsilon)$, since $|w| < 1/2$ and $|e^w - 1| \leqslant 2|w| < 2\varepsilon$.

Now since U is harmonic in $B(0, \varepsilon)$, we may write $U(w) = \mathrm{Re} f(w)$ for a function f which is analytic in $B(0, \varepsilon)$ (see Conway's classic text [18, Chapter VIII, p. 202, Theorem 2.2] for a proof). We then express f as a power series to obtain

$$U(w) = \sum_{j \geqslant 0} \mathrm{Re}(a_j w^j). \tag{3.6}$$

We now write $w = \rho e^{i\theta}$, for sufficiently small $\rho > 0$, and use the fact that $U(w) = U(\overline{w})$ to obtain

$$0 = U(w) - U(\overline{w}) = \sum_{j \geqslant 0} \mathrm{Re}(a_j \rho^j (e^{ij\theta} - e^{-ij\theta})) = \sum_{j \geqslant 0} \rho^j \mathrm{Re}(2i a_j) \sin(j\theta).$$

By the uniqueness of trigonometric series, we have that $\mathrm{Re}(2ia_j) = 0$, implying that a_j is real, for all $j \geqslant 0$. So from (3.6) and the fact that $U(0) = 0$, we obtain (3.4).

To prove the second part of the claim, we simply note that

$$U_0(w) = u_0(e^w) = u(e^w) - \mu \log|e^w| = U(w) - \mu \mathrm{Re}(w).$$

So (3.4) and the fact that $\mu = a_1$ yields (3.5). ∎

Throughout the paper, we work with the sequence $(a_j)_{j \geqslant 1}$ rather than the cumulant sequence $(\kappa_j)_{j \geqslant 1}$. We call the sequence $(a_j)_{j \geqslant 1}$ the *normalized cumulant sequence*.

3.3. *Two important "difference functions" and b-decreasing*

For $\tau \in (0, \pi)$ and $b \geqslant 0$, we define the function h_τ that compares values of u reflected about the line $\{te^{i\tau/2} : t \geqslant 0\}$

$$h_{\tau,b}(z) := u(z) - u(e^{i\tau}\overline{z}) + b.$$

We also define, for $\gamma \in (0, \pi)$, the closely related function

$$\varphi_{\gamma,b}(z) := u(z) - u(e^{i\gamma}z) + b. \tag{3.7}$$

We first observe that if u is harmonic in a sector, then $h_{\tau,b}$ and $\varphi_{\gamma,b}$ are harmonic in a slightly smaller sector for appropriately chosen τ.

Lemma 3.2. *For $R \in (1, \infty]$, $\delta > 0$, let u be a harmonic function in*[8] *$S_R(\delta)$. Then for any $\tau, \gamma \in (0, \delta/2)$, we have that the functions $h_{\tau,b}$ and $\varphi_{\gamma,b}$ are harmonic in the sector $S_R(\delta/2)$.*

Proof. We note that if v is harmonic in Ω and $\beta \in \mathbb{C} \setminus \{0\}$, then $v(\beta z)$ is harmonic in $\beta^{-1}\Omega$ and $v(\overline{z})$ is harmonic in $\{\overline{z} : z \in \Omega\}$. Also if v_1, v_2 are harmonic in Ω_1, Ω_2, respectively, then $v_1 - v_2$ is harmonic in $\Omega_1 \cap \Omega_2$ (see [4, Chapter 1]). Thus, $u(\overline{z})$ is harmonic in $S_R(\delta)$ and $u(\alpha\overline{z})$ is harmonic in $e^{i\tau}S_R(\delta) = S_R(-\delta + \tau, \delta + \tau)$ and therefore $h_{\tau,b}$ is harmonic in $S_R(-\delta + \tau, \delta + \tau) \cap S_R(\delta) \supseteq S_R(\delta/2)$. Likewise, $\varphi_{\gamma,b}$ is harmonic in $S_R(\delta - \gamma, \delta + \gamma) \cap S_R(\delta) \supseteq S_R(\delta/2)$. ∎

We now arrive at an essential definition, which we have already mentioned in the overview of the proof: b-decreasing. For $b \geqslant 0$ and $\Omega \subseteq \mathbb{C}$, we say a function u is *b-decreasing* in Ω if

$$u(\rho e^{i\theta_1}) - u(\rho e^{i\theta_2}) + b \geqslant 0$$

for all $0 \leqslant \theta_1 \leqslant \theta_2 \leqslant \pi$ with $\rho e^{i\theta_1}, \rho e^{i\theta_2} \in \Omega$. One nice feature of this definition is that u is b-decreasing in Ω if and only if u_0 is, since

$$u_0(\rho e^{i\theta_1}) - u_0(\rho e^{i\theta_2}) + b = u(\rho e^{i\theta_1}) - u(\rho e^{i\theta_2}) + b.$$

The main motivation behind this definition is easy: it is the correct definition to guarantee that the functions $h_{\tau,b}$ and $\varphi_{\gamma,b}$ are positive, for all reasonable choices of τ and γ. Later, we shall make heavy use of this fact by way of the so-called Harnack inequalities. These inequalities guarantee that h and φ do not vary too much on a set Ω away from its boundary.

Lemma 3.3. *For $\delta > 0$ and $b \geqslant 0$, let $\tau, \gamma \in (0, \delta/2)$, and let u be b-decreasing and symmetric in $S_R(\delta)$. Then $h_{\tau,b}(z) \geqslant 0$ for all $z \in S_R(\tau/2)$ and $\varphi_{\gamma,b}(z) \geqslant 0$ for all $z \in S_R(-\gamma/2, \delta - \gamma)$.*

[8] Here we understand $S_\infty(\delta)$ to mean the sector $S(\delta)$.

Proof. Let $z = re^{i\theta} \in S_R(\tau/2)$ with $|\theta| \leqslant \tau/2$. Then

$$h_{\tau,b}(re^{i\theta}) = u(re^{i\theta}) - u(re^{i(\tau-\theta)}) + b.$$

In the case $\theta \in [0, \tau/2]$, we have that $\theta \leqslant \tau - \theta$, and thus non-negativity of $h_{\tau,b}(re^{i\theta})$ follows from the b-decreasing assumption on u. On the other hand, if $\theta \in [-\tau/2, 0]$, we use symmetry to write

$$h_{\tau,b}(re^{i\theta}) = u(re^{i\theta}) - u(re^{i(\tau-\theta)}) + b = u(re^{-i\theta}) - u(re^{i(\tau-\theta)}) + b$$

with $0 \leqslant -\theta \leqslant \tau - \theta \leqslant 2\tau \leqslant \delta$; non-negativity again follows from b-decreasing. The proof for $\varphi_{\gamma,b}$ is similar. ∎

It will be important for us that Lemmas 3.2 and 3.3 tell us that both $h_{\gamma,b}$ and $\varphi_{\gamma,b}$ are harmonic and positive in a sector that contains the positive real axis with room to spare on both sides. This means that we can work near $1 \in \mathbb{C}$, without getting too close to the boundary.

4. Weakly positive and harmonic implies *b*-decreasing

In this section, we prove Lemma 4.1, our main tool for showing that a function is b-decreasing.

Lemma 4.1. *For $\delta \in (0, \pi)$ and $R > r > 0$, the following hold. Let u be a weakly positive, symmetric, harmonic function on a neighborhood of $S_R(\delta)$. If $b \geqslant 0$ is such that*

$$\left(\frac{r}{R}\right)^{1/\delta} \max_{z \in S_R^*(\delta)} |u(z)| \leqslant \frac{3b}{8}, \tag{4.1}$$

then u is b-decreasing on $S_r(\delta/2)$.

To prove Lemma 4.1, we use a well-known connection between harmonic functions and Brownian motion. The following theorem is a special case of Theorem 3.12 from the book of Mörters and Peres [54] and shows how Brownian motion can be used to recover a harmonic function from its boundary values.

Theorem 4.2 ([54, Theorem 3.12]). *Let v be a function which is harmonic on a bounded, convex set $\Omega \subseteq \mathbb{C}$ and continuous on $\partial\Omega$, let $z \in \Omega$, and let $(B_t)_{t \geqslant 0}$ be a Brownian motion started at z. If we define the stopping time*

$$\tau := \min\{t : B_t \in \partial\Omega\},$$

then we have

$$v(z) = \mathbb{E}v(B_\tau).$$

In what follows, we will understand τ to be the stopping time of a Brownian motion hitting the boundary of Ω, $\tau := \min\{t : B_t \in \partial\Omega\}$, unless otherwise stated.

4.1. *A calculation for Lemma 4.1*

To prove Lemma 4.1, we need an estimate on the probability that a Brownian motion hits one of the ends of a sector $S_R(\delta)$ before hitting the sides.

Lemma 4.3. *For $\delta \in (0, \pi)$ and $R > r > 0$, let $z \in S_r(\delta)$, and let $(B_t)_{t \geqslant 0}$ be a Brownian motion started at z and stopped when it hits $\partial S_R(\delta)$. We have*

$$\mathbb{P}(B_\tau \in S_R^*(\delta)) \leqslant \frac{4}{3}\Big(\frac{r}{R}\Big)^{c/\delta}, \tag{4.2}$$

where $c = \log 4/3$.

We should mention that there is an exact formula for the probability in (4.2) and a proof of this can be found, e.g., in [54, Theorem 7.25]. Here we include a simple proof of this weaker result, which is still sharp up to the constant c, to give the reader an intuitive feel for why we have the exponential dependence in (4.2) on δ. This dependence is quite important and ultimately explains the logarithmic factor that appears in Theorem 1.2 and Corollary 1.3.

Our first step towards Lemma 4.3 is to study a similar situation in a square. We shall then extend this to rectangles, and then use the conformal invariance of Brownian motion to finish the proof for sectors.

Observation 4.4. *For $\delta > 0$ and $y \in [-\delta, \delta]$, let E_y be the event that a Brownian motion, started at $iy \in \mathbb{C}$, hits either the left of right edges of the square $S := \{z : \mathrm{Re}(z) \leqslant \delta, \mathrm{Im}(z) \leqslant \delta\}$ before the top or bottom edges. Then $\mathbb{P}(E_y) \leqslant 3/4$.*

Proof. First, for $y \in [0, \delta]$, we consider the event H_y that a Brownian motion, started at iy, hits the top edge of S before hitting any other edge. We claim $\mathbb{P}(H_y) \geqslant 1/4$. If $y = 0$, the result is clear by symmetry. If $y > 0$, then we couple the Brownian motion $(B_t)_t$ started at 0 with a Brownian motion $(B_t + iy)_t$ started at $iy \in S$: clearly $B_t + iy$ will hit the top edge of S on every trajectory that B_t does. So $\mathbb{P}(H_y) \geqslant \mathbb{P}(H_0) = 1/4$. Now, turning to E_y, simply note that if $y \geqslant 0$, then $\mathbb{P}(E_y) \leqslant 1 - \mathbb{P}(H_y) \leqslant 3/4$. The case $y < 0$ follows by symmetry. ∎

It is now easy to deduce a version of Lemma 4.3 for rectangles. Here we see quite naturally where the exponential dependence on δ appears.

Lemma 4.5. *For $\delta > 0$ and $b > a > 0$, let $Q := \{z : |\mathrm{Re}(z)| < b, |\mathrm{Im}(z)| < \delta\}$, let $z \in Q$, and let $(B_t)_{t \geqslant 0}$ be a Brownian motion started at z which is stopped when it hits ∂Q. We have*

$$\mathbb{P}(B_\tau \in Q^*) \leqslant \exp(-c\lfloor \delta^{-1}(b-a) \rfloor),$$

where $Q^* = \{z : |\mathrm{Re}(z)| = b, |\mathrm{Im}(z)| < \delta\}$ and $c = \log 4/3$.

Proof. Let E_y be the event defined in Observation 4.4, and let $S(z')$ be the event that a Brownian motion, started at $z' \in Q$, hits one of the lines $\{z' + \delta + it\}_{t \in \mathbb{R}}, \{z' - \delta + it\}_{t \in \mathbb{R}}$ before hitting the top or bottom of Q. Clearly, $\mathbb{P}(S(z')) = \mathbb{P}(E_y)$, where $z' = x + iy$.

We now connect the rectangle-crossing to the box-crossings; a path that hits one of the ends before hitting the top or bottom of Q must cross at least $\ell \geqslant \lfloor (b-a)/\delta \rfloor$ boxes without hitting the top or bottom on Q. Therefore, we have

$$\mathbb{P}(B_\tau \in Q^*) \leqslant \sup_{z_1,\dots,z_\ell} \mathbb{P}(S(z_1) \cap \cdots \cap S(z_\ell)) \leqslant (\sup_y \mathbb{P}(E_y))^\ell, \tag{4.3}$$

where the supremum is over all complex numbers contained in Q satisfying $z_1 = z$ and $|\mathrm{Re}(z_i) - \mathrm{Re}(z_{i+1})| = \delta$, and the second inequality in (4.3) follows from the Markov property of Brownian motion. Finally, the result follows by applying Observation 4.4. ■

To prove Lemma 4.3, we simply use an analytic map to transform the rectangle Q into the truncated sector $S_R(\delta)$. The conformal invariance of Brownian motion allows us to finish. To state this property of Brownian motion a little more carefully, let $\phi\colon \mathbb{C} \to \mathbb{C}$ be an analytic function, and let $(W_t)_t$ be a Brownian motion in $\mathbb{C}$. The conformal invariance of Brownian motion means that $\phi(W_t)$ traces the path of a Brownian motion, at (possibly) different speed. In other words, there exist an increasing function $\gamma\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$, a Brownian motion (B_t) and a coupling of B_t and W_t for which $\phi(W_t) = B_{\gamma(t)}$. See the book of Mörters and Peres [54, Theorem 7.20] for a proof.

Proof of Lemma 4.3. Set $b = \log(R)$, $a = \log(r)$ and observe that the analytic function $\phi(z) := e^z$ maps the rectangle $Q = \{z : |\mathrm{Re}(z)| < b, |\mathrm{Im}(z)| < \delta\}$ to the truncated sector $S_{e^b}(\delta) = S_R(\delta)$; maps $Q = \{z : |\mathrm{Re}(z)| < a, |\mathrm{Im}(z)| < \delta\}$ to $S_r(\delta)$; and maps the ends $R^* = \{z : |\mathrm{Re}(z)| = b, |\mathrm{Im}(z)| < \delta\}$ to the ends $S^*_R(\delta)$.

To finish, choose $w \in \{s : |\mathrm{Im}(s)| < \delta, |\mathrm{Re}(s)| < b\}$ such that $\phi(w) = z$. Let $(W_t)_{t\geqslant 0}$ be a Brownian motion started at $w \in R$, let $(B_t)_t$ be a Brownian motion started at $z \in S_R(\delta)$, and let τ' be the stopping time $\tau' := \min\{t : W_t \in \partial Q\}$. By conformal invariance, there is a coupling of $(B_t)_{t\geqslant 0}$ and $(\phi(W_t))_t$ so that they trace the same path. It follows that

$$\begin{aligned}\mathbb{P}(B_\tau \in S^*_R(\delta)) &= \mathbb{P}(\phi(W_{\tau'}) \in S^*_R(\delta)) = \mathbb{P}(W_{\tau'} \in Q^*)\\ &\leqslant \exp\Big(-\Big(\log\frac{4}{3}\Big)\Big\lfloor \delta^{-1}\log\frac{R}{r}\Big\rfloor\Big),\end{aligned}$$

where the inequality follows from an application of Lemma 4.5. Utilizing $\lfloor x \rfloor \geqslant x - 1$ completes the proof. ■

4.2. *The proof of Lemma 4.1*

We now turn to the heart of Section 4, Lemma 4.6.

Lemma 4.6. *For $\delta > 0$, $R > 0$, put $\alpha = e^{i\delta}$ and let u be a weakly positive harmonic function on a neighborhood of $S_R(0,\delta)$, let $z \in S_R(0,\delta/2)$ and let $(B_t)_{t\geqslant 0}$ be a Brownian motion started at z and stopped when it hits $\partial S_R(0,\delta/2)$, then*

$$u(z) - u(\alpha\overline{z}) \geqslant -2\mathbb{P}\Big(B_\tau \in S^*_R\Big(0,\frac{\delta}{2}\Big)\Big) \max_{z\in S^*_R(\delta)} |u(z)|.$$

Proof. We define two coupled Brownian motions starting at z and $z^\circ := \alpha\overline{z}$, respectively. First, let (B_t) be a Brownian motion started at z and, in preparation for defining our Brownian motion started at z°, we define two stopping times

$$\tau = \tau_1 := \min\{t : B_t \in \partial S_R(0,\delta)\} \quad \text{and} \quad \tau_2 := \min\Big\{t : B_t \in \partial S_R\Big(0, \frac{\delta}{2}\Big)\Big\}.$$

Now, define the path $(B_t^\circ)_{t\geqslant 0}$ by $B_t^\circ := \alpha\overline{B}_t$ for $t \leqslant \tau_2$ and then $B_t^\circ := B_t$ for $t \geqslant \tau_2$. We now note that $(B_t^\circ)_t$ is in fact a Brownian motion started at z°; this is because it is a Brownian motion, by definition, after time τ_2 and it is a reflection of a Brownian motion before time τ_2, which is a Brownian motion. The only thing to note is that $B_{\tau_2} = \alpha\overline{B}_{\tau_2}$, that is, the two trajectories agree at τ_2, and thus the whole trajectory is a Brownian motion by the strong Markov property. Also note that $\tau = \min\{t : B_t^\circ \in \partial B_R(0,\delta)\}$, by symmetry.

We now apply Theorem 4.2 to u, and z, z° in the region $S(0,\delta)$ to express

$$u(z) = \mathbb{E}u(B_{\tau_1}) \quad \text{and} \quad u(z^\circ) = \mathbb{E}u(B_{\tau_1}^\circ)$$

and therefore,

$$u(z) - u(z^\circ) = \mathbb{E}(u(B_\tau) - u(B_\tau^\circ)). \tag{4.4}$$

To evaluate this expectation, we break up the space of trajectories into three events:

(1) $E_1 := \{\arg(B_{\tau_2}) = \delta/2\}$, the event that B_t, B_t° meet;

(2) $E_2 := \{\arg(B_{\tau_2}) = 0\}$, the event that B_t hits $\mathbb{R}_{\geqslant 0}$, before meeting its reflection;

(3) $E_3 := \{B_{\tau_2} \in S_R^*(\delta/2)\}$, the event that B_t hits one of the ends of the sectors, before meeting its reflection.

In the event of E_1, we have that B_t, B_t° meet before time τ, and therefore $u(B_\tau) = u(B_\tau^\circ)$ so

$$I_1 := \mathbb{E}(u(B_\tau) - u(B_\tau^\circ))\mathbb{1}(E_1) = 0. \tag{4.5}$$

In the event of E_2, $B_\tau \in \mathbb{R}_{\geqslant 0}$ and thus $B_\tau^\circ = \alpha B_\tau$ so $u(B_\tau) - u(B_\tau^\circ) \geqslant 0$, by weak positivity. In particular,

$$I_2 := \mathbb{E}(u(B_\tau) - u(\alpha B_\tau))\mathbb{1}(E_2) \geqslant 0. \tag{4.6}$$

In the case of E_3, we crudely estimate

$$I_3 := \mathbb{E}(u(B_\tau) - u(B_\tau^\circ))\mathbb{1}(E_3) \geqslant -2\mathbb{P}(E_3) \max_{z\in S_R^*(\delta)} |u(z)|. \tag{4.7}$$

Now, from (4.4), and (4.5), (4.6), (4.7), we have

$$u(z) - u(z^\circ) = I_1 + I_2 + I_3 \geqslant -2 \max_{z\in S_R^*(\delta)} |u(z)| \mathbb{P}_z\Big(B_\tau \in S_R^*\Big(0, \frac{\delta}{2}\Big)\Big),$$

as desired. ∎

We are now able to prove Lemma 4.1.

Proof of Lemma 4.1. To show that u is b-decreasing on $S_r(\delta/2)$, we let $\rho \in [1/r, r]$ and let $\theta_1, \theta_2 \in (0, \delta/2)$ satisfy $\theta_2 > \theta_1$. We need to show that

$$u(\rho e^{i\theta_1}) - u(\rho e^{i\theta_2}) \geqslant -b.$$

For this, let us put $\phi = \theta_1 + \theta_2$ and note that $\phi < \delta$. Set $z = \rho e^{i\theta_1}$ and $\alpha = e^{i\phi}$ and note that $\alpha\overline{z} = \rho e^{i\theta_2}$. We now apply Lemma 4.6 with $\delta = \phi$ to obtain

$$\begin{aligned} u(\rho e^{i\theta_1}) - u(\rho e^{i\theta_2}) &\geqslant -2\mathbb{P}\Big(B_\tau \in S_R^*\Big(0, \frac{\phi}{2}\Big)\Big) \max_{z \in S_R^*(\phi)} |u(z)| \\ &\geqslant -2 \cdot \frac{4}{3}\Big(\frac{r}{R}\Big)^{4c/\delta} \cdot \max_{z \in S_R^*(\delta)} |u(z)| \\ &\geqslant -b, \end{aligned}$$

where the penultimate inequality holds due to the fact that $S_R^*(0, \phi/2)$ is a sector of width at most $\delta/2$ and so we apply Lemma 4.3 to the sector $S_R(\delta/4)$. The last inequality holds by the condition on b and the inequality $4\log(4/3) > 1$. Hence we have shown that u is b-decreasing in $S_r(0, \delta/2)$. ∎

5. A key comparison

With our main positivity hypothesis in place, we now start with the first in a series of steps to prove Lemma 8.1. In Sections 5, 6 and 7, we build up the ingredients for the proof of Lemma 8.1, finally stated and proved in Section 8. The objective of this section is to prove Lemma 5.1, which says that under the hypotheses of Lemma 8.1, we can bound the maximum value of u_0 in a small box around 1 in terms of the much more amenable function $\varphi_{\gamma,b}$.

Lemma 5.1. *For $b \geqslant 0$, $\varepsilon \in (0, 1/8)$ and $\eta \in (0, \varepsilon]$, let $u(z)$ be a b-decreasing, weakly positive, symmetric and harmonic function on $B(1, 8\varepsilon)$, for which $u(1) = 0$. Let u_0 and $\varphi_{\eta,b}$ be the associated functions defined at (3.3) and (3.7). We have that*

$$\max_{z \in B(1,\varepsilon)} |u_0(z)| \leqslant 3^4 \cdot 3^{96\varepsilon/\eta} \varphi_{\eta,b}(1).$$

To prove this lemma, we make a few preparations.

5.1. Positive on the real line

In our first step towards Lemma 5.1, we show that the function

$$u_0(z) = u(z) - \mu \log|z|$$

is positive on the positive real axis. To prove this, we first need the following basic fact, which first appears in the work of De Angelis [21] and then was slightly[9] extended in the work of Bergweiler, Eremenko and Sokal [7, 8]. We include a short proof.

[9] De Angelis actually assumes that $u = \log|p|$ for a polynomial p.

Lemma 5.2. *Let $r > 0$. If u is weakly positive, symmetric and harmonic on a neighborhood of r, then $U''(\log r) \geqslant 0$.*

Proof. Write $z = x + iy$ and put $V(x, y) = U(x + iy)$. Note that the harmonicity of U at $a := \log r$ implies $V_{xx}(a, 0) + V_{yy}(a, 0) = 0$. Weak positivity and symmetry of u implies that $V_{yy}(a, 0) \leqslant 0$, since $V_{yy}(a, 0)$ can be written

$$\lim_{h\to 0} \frac{V(a,h) + V(a,-h) - 2V(a,0)}{h^2} = \lim_{h\to 0} \frac{u(re^{ih}) + u(re^{-ih}) - 2u(r)}{h^2} \leqslant 0.$$

Thus,

$$U''(a) = V_{xx}(a, 0) \geqslant 0.$$ ∎

Remark 5.3. In the case $u = \log|f_X|$, there is a natural probabilistic interpretation of Lemma 5.2 that can be turned into a proof. After unwinding the definitions a little, one can see that Lemma 5.2 simply says that the random variable defined by the probability generating function $f_X(rz) f_X^{-1}(r)$, has *non-negative* variance. This is trivially true, as all random variables have non-negative variance. However, we have elected to include this more general result, as it will simplify our exposition.

We now deduce the following small but crucial ingredient in the proof of Lemma 5.1: u_0 is non-negative on the positive real axis.

Lemma 5.4. *For $\varepsilon \in (0, 1)$, let $E \subseteq \mathbb{C}$ be an open set containing the interval $[1-\varepsilon, 1+\varepsilon]$. If u is weakly positive, harmonic and symmetric in E, then $u_0(r) \geqslant u(1)$ for all $r \in (1-\varepsilon, 1+\varepsilon)$.*

Proof. We may write $r = e^t$ for some $t \in \mathbb{R}$ and apply Taylor's theorem to $U(t)$ at $t = 0$ to obtain

$$U(t) - U(0) - tU'(0) = \frac{t_0^2}{2} U''(t_0),$$

for some t_0 with $t_0 \in (1-\varepsilon, 1+\varepsilon)$. Since $U_0(t) = U(t) - tU'(0)$ ((3.5) in Lemma 3.1), we have

$$U_0(t) - U(0) = \frac{t_0^2}{2} U''(t_0) \geqslant 0,$$

by Lemma 5.2 and so

$$u_0(r) = u_0(e^t) = U_0(t) \geqslant u(1),$$

as desired. ∎

5.2. The Poisson density and Harnack inequalities

In Section 4, we saw that we could use Brownian motion to recover the values of a harmonic function on Ω from the values of its boundary $\partial\Omega$, by using Theorem 4.2. In particular, we had that

$$u(z) = \mathbb{E}u(B_\tau), \tag{5.1}$$

where $(B_t)_{t\geqslant 0}$ is a Brownian motion started at $z \in \Omega$ and τ is the stopping time of hitting $\partial\Omega$. In the case that Ω is a ball B, the expectation in (5.1) has a density function P_z (with respect to the Lebesgue measure on the circle), and so we can express

$$v(z) = \int_{\partial\Omega} v(s) P_z(s)\, \mathrm{d}s.$$

We define the collection of functions $\{P_z\}$ as the *Poisson density* of the circle[10] and note that P_z are non-negative functions of the boundary. We refer the reader to [4, 18] for a general treatment of harmonic functions on a ball.

We now state an important tool for working with positive harmonic functions, the *Harnack inequalities* for the ball. These say that if v is a positive harmonic function on the ball $B(0, 2\varepsilon)$ and if $z \in B(0, \varepsilon) \subseteq B(0, 2\varepsilon)$, then

$$\frac{1}{3} \leqslant \frac{v(z)}{v(0)} \leqslant 3. \tag{5.2}$$

A statement and proof of this result can be found in [4, p. 47]. For our purposes, we need a slight generalization.

Lemma 5.5. *Let v be a positive harmonic function on an open set $\Omega \subset \mathbb{C}$. For $\varepsilon > 0$, let $z_1, z_2 \in \Omega$ be points at distance $d := d(z_1, z_2)$ so that all z_3 that lie on the line segment joining z_1 and z_2 have $B(z_3, \varepsilon) \subset \Omega$. Then*

$$\frac{1}{3^{2d/\varepsilon+1}} \leqslant \frac{v(z_1)}{v(z_2)} \leqslant 3^{2d/\varepsilon+1}.$$

We will also need the following lemma, which is a simple consequence of the Harnack inequalities.

Lemma 5.6. *For $\varepsilon > 0$, let $\{P_z\}$ be the Poisson density of the ball $B(1, 2\varepsilon)$. Then for all $z \in B(1, \varepsilon)$, we have*

$$\max_{s \in \partial B_{2\varepsilon}} \frac{P_z(s)}{P_1(s)} \leqslant 3.$$

Both Lemmas 5.5 and 5.6 follow easily from (5.2). We now prove a straightforward lemma that will allow us to find a large negative value of u_0. For this lemma, we note that if $\{P_z\}$ is the Poisson density of $B(1, \varepsilon)$ then, by symmetry, $P_1(s) = 1/(2\pi\varepsilon)$ for all $s \in \partial B(1, \varepsilon)$.

Lemma 5.7. *Let $\varepsilon > 0$, and let $v(z)$ be a symmetric, harmonic function on $B(1, \varepsilon)$ with $v(1) = 0$, let $\{P_z\}$ be the Poisson density of the ball $B(1, \varepsilon)$ and set*

$$M := \int_{s \in \partial B(1,\varepsilon)} |v(s)| P_1(s)\, \mathrm{d}s.$$

Then there is a $z_0 \in \partial B(1, \varepsilon)$ with $\mathrm{Im}(z_0) \geqslant 0$ and $v(z_0) \leqslant -M/2$.

[10] This definition is designed to parallel the (standard) definition of the *Poisson kernel* which is defined with respect to the complex line integral, rather than the uniform measure on the circle.

Proof. Let us set $B_\varepsilon := B(1,\varepsilon)$, $P(s) := P_1(s)$ and define $\partial B_\varepsilon = E_+ \cup E_-$, where $E_+ = \{s \in \partial B_\varepsilon : v(s) \geqslant 0\}$ and $E_+ = \{s \in \partial B_\varepsilon : v(s) \leqslant 0\}$. Now put

$$A = \int_{s\in E+} |v(s)|P(s)\,\mathrm{d}s \quad \text{and} \quad B = \int_{s\in E-} |v(s)|P(s)\,\mathrm{d}s$$

and note that

$$0 = v(1) = \int_{s\in\partial B_\varepsilon} v(s)P(s)\,\mathrm{d}s = \int_{s\in E+} |v(s)|P(s)\,\mathrm{d}s - \int_{s\in E-} |v(s)|P(s)\,\mathrm{d}s$$
$$= A - B$$

and that $A + B = M$. Thus,

$$\frac{M}{2} = \int_{E-} |v(s)|P(s)\,\mathrm{d}s \leqslant \max_{s\in E-} |v(s)| \int_{E-} P(s)\,\mathrm{d}s \leqslant \max_{s\in E-} |v(s)|,$$

where the last line holds due to the facts that $P(s) = 1/(2\pi\varepsilon)$ and the length of E^- is at most $2\pi\varepsilon$. So if we let z_0 be a value which attains this maximum, we note that both $v(z_0) = v(\overline{z}_0) \leqslant -M/2$ by symmetry. Hence one of z_0, $\overline{z}_0$ will have non-negative imaginary part, as desired. ∎

5.3. *Proof of Lemma* 5.1

We now are in a position to prove the main result of this section, Lemma 5.1.

Proof of Lemma 5.1. To reduce clutter, let us define B_ε, $B_{2\varepsilon}$ to be $B(1,\varepsilon)$, $B(1,2\varepsilon)$, respectively. Let P_z be the Poisson density of $B_{2\varepsilon}$ and put

$$M := \int_{s\in\partial B_{2\varepsilon}} |u_0(s)|P_1(s)\,\mathrm{d}s.$$

By the definition of the Poisson density, for each $z \in B_\varepsilon$, we have

$$u_0(z) = \int_{s\in\partial B_{2\varepsilon}} u_0(s)P_z(s)\,\mathrm{d}s,$$

and since, for all $z \in B_\varepsilon$, we have $\max_{s\in\partial B_{2\varepsilon}} P_z(s)/P_1(s) \leqslant 3$ (by Lemma 5.6), we obtain

$$\max_{z\in B_\varepsilon} |u_0(z)| \leqslant 3 \int_{s\in\partial B_{2\varepsilon}} |u_0(z)|P_1(s)\,\mathrm{d}s = 3M. \tag{5.3}$$

We now apply Lemma 5.7 to the function $v := u_0(z)$ (which is harmonic, symmetric and $u_0(1) = 0$) and the region $B_{2\varepsilon}$ to find a point $z_0 \in \partial B_{2\varepsilon}$ for which $u_0(z_0) \leqslant -M/2$ and $\mathrm{Im}(z_0) \geqslant 0$. We may write z_0 in the form $z_0 = \rho e^{i\phi}$, where $\phi \in [0, 4\varepsilon]$ and $\rho \in [1-2\varepsilon, 1+2\varepsilon]$.

Further, $\rho e^{4\varepsilon i} \in B(1, 8\varepsilon)$, allowing us to make use of the b-decreasing hypothesis:

$$u_0(\rho e^{4\varepsilon i}) \leqslant u_0(z_0) + b \leqslant b - \frac{M}{2}.$$

We now apply Lemma 5.4 to see that $u_0(\rho) \geqslant 0$ as $\rho > 0$, thereby allowing us to obtain a bound on $h_{4\varepsilon,b}(\rho)$,

$$h_{4\varepsilon,b}(\rho) = u_0(\rho) - u_0(\rho e^{i4\varepsilon}) + b \geqslant -u_0(\rho e^{i4\varepsilon}) + b \geqslant \frac{M}{2}. \tag{5.4}$$

We know that $S(-4\varepsilon, 4\varepsilon) \cap B_{4\varepsilon} \subset B_{8\varepsilon}$ and u is b-decreasing in $B_{8\varepsilon}$. We then know that $h_{4\varepsilon,b}(z)$ is harmonic and positive (Lemmas 3.2 and 3.3) in $S(-2\varepsilon, 2\varepsilon) \cap B_{4\varepsilon}$ and thus we may apply Lemma 5.5 to learn that

$$h_{4\varepsilon,b}(\rho) \leqslant 3^{64\varepsilon/\eta+1} h_{4\varepsilon,b}(\rho e^{i(2\varepsilon-\eta/2)}), \tag{5.5}$$

since the distance $d(\rho, \rho e^{i(2\varepsilon-\eta/2)}) \leqslant (1+2\varepsilon)2\varepsilon \leqslant 4\varepsilon$ and each point on the segment between them is at least $\rho \cdot \eta/4 \geqslant \eta/8$ from the boundary of $S(-2\varepsilon, 2\varepsilon) \cap B_{4\varepsilon}$.

Now observe that at the value of $z = \rho e^{i(2\varepsilon-\eta/2)}$, we have $h_{4\varepsilon,b}(z) = \varphi_{\eta,b}(z)$. That is,

$$h_{4\varepsilon,b}(\rho e^{i(2\varepsilon-\eta/2)}) = u(\rho e^{i(2\varepsilon-\eta/2)}) - u(\rho e^{i(2\varepsilon+\eta/2)}) + b = \varphi_{\eta,b}(\rho e^{i(2\varepsilon-\eta/2)}). \tag{5.6}$$

We now apply Lemmas 3.2 and 3.3 to learn that the function $\varphi_{\eta,b}$ is harmonic and positive in $S(-\eta/2, 4\varepsilon - \eta) \cap B_{4\varepsilon}$. Hence we may apply Lemma 5.5 along with the fact that $\rho \in [1-2\varepsilon, 1+2\varepsilon]$ to see that

$$\varphi_{\eta,b}(\rho e^{i(2\varepsilon-\eta)}) \leqslant 3^{32\varepsilon/\eta+1} \varphi_{\eta,b}(1), \tag{5.7}$$

since $|\rho e^{i(2\varepsilon-\eta/2)} - 1| \leqslant 4\varepsilon$ and each point on the segment between them is at least $\eta/4$ from the boundary of $S(-\eta/2, 4\varepsilon - \eta) \cap B_{4\varepsilon}$. Thus, chaining together lines (5.3), (5.4), (5.5), (5.6), (5.7) gives $\max_{z \in B_\varepsilon} |u_0(z)| \leqslant 3^4 \cdot 3^{96\varepsilon/\eta} \varphi_{\eta,b}(1)$, as desired. ∎

6. Bounding the tail of the cumulant sequence

In the previous section, we showed how to control the maximum of u_0 in a small ball around 1, in terms of the much more amenable function $\varphi_{\eta,b}$. In this section, we use this bound to prove that the normalized cumulant sequence $(a_j)_{j \geqslant 2}$ has nice decay properties.

Lemma 6.1. *For $\varepsilon \in (0, 2^{-4})$ and $b \geqslant 0$, let u be a b-decreasing weakly positive symmetric harmonic function in $B(1, 2^4\varepsilon)$. Let $(a_j)_{j \geqslant 1}$ be the normalized cumulant sequence of u. If*

$$\sum_{j \geqslant 2} |a_j| \varepsilon^j > b, \tag{6.1}$$

then for all $L \geqslant 2$ we have

$$\frac{\sum_{j \geqslant L} |a_j| \varepsilon^j}{\sum_{j \geqslant 2} |a_j| \varepsilon^j} \leqslant C \cdot 2^{-L}, \tag{6.2}$$

where $C > 0$ is an absolute constant.[11]

[11] Indeed, we can take $C = 3^{390}$.

Proof. We first point out that the expression in (6.2) makes sense; the denominator is non-zero from the strict inequality at (6.1), and the numerator is finite since we may write

$$U_0(w) = \sum_{j \geqslant 2} a_j \mathrm{Re}(w^j), \tag{6.3}$$

by Lemma 3.1, for all $w \in B(0, 8\varepsilon)$ and thus the series $\sum a_j \varepsilon^j$ is absolutely convergent.

To prove Lemma 6.1, the idea is to compare both the numerator and denominator in (6.2) to $\varphi_{\varepsilon,b}(1)$. We begin with the denominator. Recalling that

$$\varphi_{\varepsilon,b}(e^w) = u(e^w) - u(e^{w+i\varepsilon}) + b,$$

we use (6.3) to express

$$\varphi_{\varepsilon,b}(e^w) = \sum_{j \geqslant 2} a_j \mathrm{Re}(w^j - (w + i\varepsilon)^j) + b$$

for w sufficiently small. And so, setting $w = 0$, we obtain

$$|\varphi_{\varepsilon,b}(1)| \leqslant \sum_{j \geqslant 2} |a_j|\varepsilon^j + b \leqslant 2 \sum_{j \geqslant 2} |a_j|\varepsilon^j \tag{6.4}$$

by the triangle inequality and our assumption at (6.1).

We now turn to obtain an upper bound on the numerator of (6.2). We apply the Cauchy–Schwarz inequality to obtain

$$\sum_{j \geqslant L} |a_j|\varepsilon^j \leqslant 2\Big(\sum_{j \geqslant 2} |a_j|^2 (2\varepsilon)^{2j}\Big)^{1/2} 2^{-L}. \tag{6.5}$$

We now look to relate the series on the right-hand side of (6.5) to U_0. In preparation for this, we write

$$U_0(\rho, \theta) := U_0(\rho e^{i\theta}) = \sum_{j \geqslant 2} a_j \rho^j \cos(j\theta),$$

which is valid for all $|\rho| < 8\varepsilon$, due to (6.3), and then use Parseval's theorem to write

$$\sum_{j \geqslant 2} |a_j|^2 (2\varepsilon)^{2j} = \frac{1}{2\pi} \int_0^{2\pi} |U_0(2\varepsilon, \theta)|^2 \mathrm{d}\theta \leqslant \max_{\theta \in [0, 2\pi]} |U_0(2\varepsilon, \theta)|^2. \tag{6.6}$$

As a remark, note that (6.6) along with (6.1) implies that $\max_{\theta \in [0,2\pi]} |U_0(2\varepsilon, \theta)| > 0$.

Returning to the main arc of the proof, recall that $z = e^w$ and $w = \rho e^{i\theta}$; so as θ ranges over $[0, 2\pi]$, z lies on the curve

$$\Gamma = \{\exp(2\varepsilon e^{i\theta}) : \theta \in [0, 2\pi]\},$$

which is contained in the ball $B(1, 4\varepsilon)$, due to the inequality $|1 - \exp(2\varepsilon e^{i\theta})| \leqslant 4\varepsilon$, which holds for $\varepsilon < 1$. Hence we may bound the right-hand side of (6.6)

$$\max_{\theta \in [0,2\pi]} |U_0(2\varepsilon, \theta)|^2 \leqslant \max_{z \in B(1,4\varepsilon)} |u_0(z)|^2. \tag{6.7}$$

Here is the key ingredient: we apply Lemma 5.1 to obtain an upper bound on u_0 in terms of $\varphi_{\eta,b}$ in $B(1,4\varepsilon)$ with $\eta = \varepsilon$,

$$\max_{z\in B(1,4\varepsilon)} |u_0(z)|^2 \leqslant (3^{388}\varphi_{\varepsilon,b}(1))^2. \tag{6.8}$$

Note that this also implies that $\varphi_{\varepsilon,b}(1) > 0$, due to the remark after (6.6) and (6.7), (6.8).

To finish, we put together the lower bound at (6.4) on the denominator in (6.2) with the upper bound on the numerator, coming from (6.5), (6.6), (6.7) and (6.8), to obtain

$$\frac{\sum_{j\geqslant L} |a_j|\varepsilon^j}{\sum_{j\geqslant 2} |a_j|\varepsilon^i} \leqslant 2\cdot\frac{3^{388}\varphi_{\varepsilon,b}(1)}{\varphi_{\varepsilon,b}(1)/2}2^{-L} \leqslant 3^{390}2^{-L},$$

as desired. ∎

7. Taming the cumulant sequence

In this section, we provide a third and final ingredient in our proof of Lemma 8.1, our core technical lemma. In Section 6, we showed that the sequence $(a_j)_{j\geqslant 2}$ had to have quite a bit of its "mass" concentrated on the early terms. In this section, we use our weak positivity hypothesis to show that, in this situation, we can control all of the cumulants in terms of the *second* cumulant, the variance.

The main result of this section is Lemma 7.5, which can be seen as a quantitative version of a tool co-discovered by De Angelis [21] and Bergweiler, Eremenko and Sokal [8] which was used in their work on classifying polynomials whose large powers have all positive coefficients. It is also a relative of Lemma 7 in the previous work of the authors [51] and can be viewed as an effective form of Marcinkiewicz's theorem [48].

To prove Lemma 7.5, we need the following preparatory lemma, which is an elementary fact about sequences of non-negative real numbers.

Lemma 7.1. *Let $A \geqslant 1$, $s > 0$, and let $(c_i)_{i\geqslant 1}$ be a sequence of non-negative numbers for which the sum $\sum_{i\geqslant 1} c_i s^i$ converges and is non-zero. If $L \in \mathbb{N}$ is such that*

$$\sum_{i=1}^{L} c_i s^i > \sum_{i>L} c_i s^i,$$

then there exist an $\ell \in \{1,\dots,L\}$ and $s_* > s(16A)^{-(L+1)}$ such that

$$c_\ell s_*^\ell > A\sum_{i\neq\ell} c_i s_*^i. \tag{7.1}$$

Proof. To start, we choose $s_0 := s/(2A)$. This immediately gives us

$$\sum_{i=1}^{L} c_i s_0^i > 2A\sum_{i>L} c_i s_0^i. \tag{7.2}$$

We now define an algorithm that will find $\ell \in \{1, \ldots, L\}$ and $s_* > s(16A)^{-(L+1)}$ that satisfy (7.1). Initialize $t = 0$, $s_0 = s_0$ (defined above) and $j_0 = L$ and inductively define a sequence of integers $j_1 \geqslant j_2 \geqslant \cdots \geqslant 1$ and positive real numbers $s_1 > s_2 > \cdots$ as follows: if the pair (j_t, s_t) satisfies

$$c_{j_t} s_t^{j_t} > 2A \sum_{\substack{1 \leqslant i \leqslant L \\ i \neq j_t}} c_i s_t^i, \tag{7.3}$$

then we terminate and return $(\ell, s_*) = (j_t, s_t)$. Otherwise, choose j_{t+1} such that

$$c_{j_{t+1}} s_t^{j_{t+1}} = \max\{c_1 s_t^1, c_2 s_t^2, \ldots, c_t s_t^{j_t}\} \tag{7.4}$$

and set $s_{t+1} = s_t/(16A)$. To see that this algorithm successfully produces a pair (ℓ, s_*) that satisfies the conclusions of the lemma, we prove two claims.

Claim 7.2. *For each $t \geqslant 0$, we have*

$$c_{j_t} s_t^{j_t} > 4A \sum_{i=j_t+1}^{L} c_i s_t^i. \tag{7.5}$$

Proof. We apply induction on t; note that the $t = 0$ case is trivial. Now suppose (7.5) is satisfied for some $t \geqslant 0$, write $a = j_t$, $b = j_{t+1}$ (for ease of notation) and recall that $b = j_{t+1}$ was chosen so that $c_b s_t^b = \max_{1 \leqslant i \leqslant a}\{c_i s_t^i\}$; thus,

$$c_b s_t^b > \sum_{i=b+1}^{a} \frac{c_i s_t^i}{2^{i-b}},$$

and since $s_{t+1} = s_t/(16A)$, we have

$$c_b s_{t+1}^b = (16A)^{-b} c_b s_t^b > (16A)^{-b} \sum_{i=b+1}^{a} \frac{c_i s_t^i}{2^{i-b}} \geqslant 8A \sum_{i=b+1}^{a} c_i s_{t+1}^i. \tag{7.6}$$

By the induction hypothesis at (7.5), we have $c_a s_t^a \geqslant 4A \sum_{i=a+1}^{L} c_i s_t^i$ and thus (crudely) we have

$$c_b s_{t+1}^b \geqslant 8A \sum_{i=a+1}^{L} c_i s_{t+1}^i. \tag{7.7}$$

Averaging (7.6) and (7.7) yields

$$c_b s_{t+1}^b > 4A \sum_{i=b+1}^{L} c_i s_{t+1}^i,$$

as desired. This completes the proof of the claim, by induction. ∎

Claim 7.3. *We have that $j_1 > j_2 > \cdots \geqslant 1$ is a strictly decreasing sequence of integers.*

Proof. By definition, we have $j_1 \geqslant j_2 \geqslant \cdots$ and so we claim that if $j_{t+1} = j_t$, then the pair (j_t, s_t) would in fact satisfy (7.3), the halting condition for the algorithm. So suppose that $j := j_{t+1} = j_t$ and recall that $s_{t+1} = s_t/(16A)$; then by (7.4), we have, for all $i \leqslant j$,

$$c_i s_t^i \leqslant c_j s_t^j (16A)^{i-j}.$$

This implies

$$4A \sum_{i=1}^{j-1} c_i s_t^i \leqslant c_j s_t^j \left(4A \sum_{i=1}^{j-1} (16A)^{i-j} \right) \leqslant c_j s_t^j. \tag{7.8}$$

Averaging (7.8) and (7.5) yields (7.3) for $(\ell, s_*) = (j_t, s_t)$, implying that the algorithm would have halted before proceeding to step $t+1$, a contradiction. ∎

Thus, Claim 7.3 tells us that the algorithm must terminate in at most L steps and hence $s_* > s(16A)^{-(L+1)}$.

To see that we have found a pair (ℓ, s_*) that also satisfies (7.1), we simply note that (7.2) implies $\sum_{i=1}^{L} c_i s_*^i > 2A \sum_{i>L} c_i s_*^i$ and thus, averaging this with (7.3), yields the inequality (7.1), as desired. ∎

For our main lemma of this section, we make use of the (somewhat crude) inequalities.

Fact 7.4. For $j \geqslant 3$, we have

$$\min_{\theta \in \mathbb{R}} \{(\cos\theta)^j - \cos j\theta\} < -\frac{1}{2}, \quad \max_{\theta \in \mathbb{R}} \{(\cos\theta)^j - \cos j\theta\} > \frac{1}{2}.$$

As mentioned before, we apply a clever idea from the works of De Angelis and Bergweiler, Eremenko and Sokal and use the non-negativity of another "difference function":

$$u(|z|) - u(z).$$

Lemma 7.5. *For $s \in (0, 1/2)$ and $L \geqslant 2$, let u be a weakly positive, symmetric harmonic function on $B(1, 2s)$, and let $(a_j)_j$ be its normalized cumulant sequence. If $(a_j)_{j\geqslant 2}$ is a non-zero sequence and*

$$\sum_{j\geqslant 2}^{L} |a_j| s^j \geqslant \sum_{j>L} |a_j| s^j,$$

then there exists a real number $s_* > s2^{-6(L+1)}$ for which

$$|a_2| \geqslant s_*^{j-2} |a_j|$$

for all $j \geqslant 2$.

Proof. First note that the function $U(w) = u(e^w)$ is harmonic for $w \in B(0, s)$ due to the inequality $|e^w - 1| \leqslant 2|w|$ for $|w| \leqslant 1/2$ and the fact that u is harmonic on $B(1, 2s)$. We consider the function

$$U_0(\mathrm{Re}(w)) - U_0(w) = u_0(|e^w|) - u_0(e^w) \geqslant 0, \tag{7.9}$$

where the inequality follows from weak positivity. Now, writing $w = \rho e^{i\theta}$ and considering the series expansion of U_0 around $w = 0$ (Lemma 3.1), we have

$$F(\rho,\theta) := U_0(\rho\cos\theta) - U_0(\rho e^{i\theta}) = \sum_{j\geqslant 2} a_j\rho^j((\cos\theta)^j - \cos j\theta)$$

for all $0 \leqslant \rho < s$. Since a_j is not identically 0 for all $j \geqslant 2$, we may apply Lemma 7.1 to the sequence $(|a_j|)_{j\geqslant 2}$ with $A = 4$, to get an integer $\ell \in [L]$ and a real number $s_* > s2^{-6(L+1)}$ such that

$$|a_\ell|s_*^\ell > 4\sum_{2\leqslant i\neq\ell}|a_i|s_*^i. \tag{7.10}$$

We now use weak positivity to see that $\ell = 2$. For this, assume $\ell > 2$ and apply Fact 7.4 to find a θ_0 for which

$$a_\ell((\cos\theta_0)^\ell - \cos\ell\theta_0) \leqslant -\frac{|a_\ell|}{2}. \tag{7.11}$$

We write

$$F(s_*,\theta_0) = a_\ell s_*^\ell((\cos\theta_0)^\ell - \cos\ell\theta_0) + \sum_{2\leqslant j\neq\ell} a_j s_*^j((\cos\theta_0)^j - \cos j\theta_0)$$

and apply (7.11) to bound the first term on the right-hand side and apply the triangle-inequality to bound the sum. We obtain

$$F(s_0,\theta_0) \leqslant \frac{-|a_\ell|s_*^\ell}{2} + 2\sum_{2\leqslant j\neq\ell}|a_j|s_*^j < 0,$$

where the last inequality follows from (7.10). However, this contradicts the positivity of F (7.9). We therefore conclude that $\ell = 2$ and so, from (7.10) again, we have that

$$|a_2|s_*^2 > 4\sum_{i\geqslant 3}|a_i|s_*^i \geqslant 4|a_j|s_*^j$$

for any $j \geqslant 3$, as desired. ∎

8. Proof of Lemma 8.1: The final stroke

In this section, we combine the ingredients from the previous sections to prove Lemma 8.1. What we state here is slightly stronger than what we need, but we make these bounds explicit for use in later work.

Lemma 8.1. *For $\varepsilon \in (0,1)$, $b \geqslant 0$ and $n \geqslant 1$, let X be a random variable with standard deviation $\sigma > 0$, logarithmic potential $u = u_X$ and normalized cumulant sequence $(a_j)_{j\geqslant 1}$. If u is b-decreasing and harmonic in $B(1,\varepsilon)$ and*

$$\sum_{j\geqslant 2}|a_j|\left(\frac{\varepsilon}{32}\right)^j > b, \tag{8.1}$$

then ψ_{X^*}, *the characteristic function of* $X^* := (X-\mu)\sigma^{-1}$, *satisfies*

$$\psi_{X^*}(\xi) = \exp\Big(-\frac{\xi^2}{2} + R(\xi)\Big),$$

where

(1) $R(0) = R^{(1)}(0) = R^{(2)}(0) = 0$ *and*

$$|R^{(\ell)}(0)| \leqslant \ell!(c_2\sigma)^{2-\ell}$$

for all $\ell \geqslant 3$.

(2) *In particular,*[12]

$$|R(\xi)| \leqslant \frac{c_1|\xi|^3}{\varepsilon\sigma}$$

for all $\xi \in \mathbb{C}$ *with* $|\xi| \leqslant c_2\varepsilon\sigma$.

Proof. Let $\psi_X(\xi) = \mathbb{E}_X e^{i\xi X}$ be the characteristic function of X, and note that

$$\psi_X(\xi) = \exp\Big(\sum_{j\geqslant 1}\frac{\kappa_j}{j!}(i\xi)^j\Big) = \exp\Big(\sum_{j\geqslant 1} a_j(i\xi)^j\Big),$$

where κ_j is the j-th cumulant of X and $(a_j)_{j\geqslant 1}$ is the normalized cumulant sequence. Here, this expansion is valid for all $|\xi| < \varepsilon/2$ since harmonicity of u in $B(1,\varepsilon)$ implies analyticity of ψ in $B(0,\varepsilon/2)$ due to the inequality $|1-e^w| \leqslant 2|w|$ for $|w| < 1/2$. Now note that $\psi_{X^*}(\xi) = \psi_X(\xi/\sigma)e^{-i\mu\xi/\sigma}$ is the characteristic function of X^*. Using the fact that $a_1 = \mu$ and $a_2 = -\sigma/2$, as noted at (3.1) and (3.2), we have

$$\psi_{X^*}(\xi) = \exp\Big(-\frac{\xi^2}{2} + \sum_{j\geqslant 3}\frac{a_j}{\sigma^j}(i\xi)^j\Big)$$

and so we define

$$R(\xi) := \sum_{j\geqslant 3}\frac{a_j}{\sigma^j}(i\xi)^j.$$

We now apply Lemma 6.1 to bound R. To see that we may apply this lemma, note that (8.1) implies condition (6.1) in Lemma 6.1; the logarithmic potential $u = u_X$ is weakly positive and symmetric in $B(1,\varepsilon)$ (as noted in Section 3); and u is b-decreasing and harmonic in $B(1,\varepsilon)$, by assumption. Therefore,

$$\frac{\sum_{j\geqslant L}|a_j|(\varepsilon/32)^j}{\sum_{j\geqslant 2}|a_j|(\varepsilon/32)^j} \leqslant C\cdot 2^{-L}$$

for all $L \geqslant 2$. Now, if we choose $L = 2 + \log_2 C$, we have that

$$\sum_{j=2}^{L}|a_j|\Big(\frac{\varepsilon}{32}\Big)^j > \sum_{j>L}|a_j|\Big(\frac{\varepsilon}{32}\Big)^j,$$

[12] We can take the constants $c_1 = 2^{3246}$, $c_2 = 2^{-3246}$.

and so we may apply Lemma 7.5 with $L = 539$ and $s = \varepsilon/32$ to obtain an $s_* > 2^{-3245}\varepsilon$ for which

$$\sigma^2 = |a_2| > s_*^{j-2}|a_j|.$$

And so for $j \geqslant 3$, the j-th term in the expansion of $R(\xi)$ is

$$\frac{|R^{(j)}(0)|}{j!} = |a_j|\sigma^{-j} \leqslant (s_*\sigma)^{2-j},$$

and so, for $|\xi| < s_*\sigma$, we have

$$|R(\xi)| \leqslant \sum_{j\geqslant 3} \frac{|a_j||\xi|^j}{\sigma^j} \leqslant \sum_{j\geqslant 3} \frac{|\xi|^j}{(s_*\sigma)^{j-2}} = \frac{|\xi|^3}{s_*\sigma(1-|\xi|/(s_*\sigma))}.$$

This means that we can factor

$$\psi_{X^*}(\xi) = e^{-\xi^2/2}e^{R(\xi)},$$

where $|R(\xi)| < 2|\xi|^3/(s_*\sigma) \leqslant 2^{3246}/(\varepsilon\sigma)$ for $|\xi| < (s_*\sigma)/2 \leqslant 2^{-3246}\varepsilon\sigma$. This completes the proof of Lemma 8.1. ∎

9. Proofs of Theorems 1.4 and 1.2

In this section, we use Lemma 4.1 along with our main technical lemma, Lemma 8.1, to deduce our theorems on univariate polynomials. Before we finish these proofs, we need to quickly derive our "Fourier inversion" lemma, which allows us to conclude that X is approximately normal based on the hypothesis that the characteristic function ψ_X is approximately the characteristic function of a normal.

9.1. Fourier inversion

In this short subsection, we derive the following basic "Fourier inversion" tool.

Lemma 9.1. *Let $X \in \mathbb{R}$ be a random variable with characteristic function ψ. If*

$$\psi(\xi) = \exp\Big(-\frac{\xi^2}{2} + R(\xi)\Big),$$

where $|R(\xi)| \leqslant \eta|\xi|^3$ for all $|\xi| < \tau$, then

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X \leqslant t) - \mathbb{P}(Z \leqslant t)| \leqslant 2^9 \max\{\eta, \tau^{-1}\},$$

where $Z \sim N(0,1)$.

Following Lebowitz, Pittel, Ruelle and Speer, we use the following quantitative result which can be found in the textbook of Feller [25, p. 538].

Lemma 9.2. *Let $Z \sim N(0,1)$ be a standard normal variable, let $X \in \mathbb{R}$ be a random variable, and let $\psi(\xi)$ be its characteristic function. Then, for all $T > 0$, we have*

$$\sup_{t \in R} |\mathbb{P}(X \leqslant t) - \mathbb{P}(Z \leqslant t)| \leqslant \frac{1}{\pi} \int_{-T}^{T} \left| \frac{\psi(\xi) - e^{-\xi^2/2}}{\xi} \right| d\xi + \frac{4}{T}.$$

We may now easily derive Lemma 9.1, our Fourier inversion lemma.

Proof of Lemma 9.1. We apply Lemma 9.2 with $T = \min\{\eta^{-1}/8, \tau\}$ to obtain

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X \leqslant t) - \mathbb{P}(Z \leqslant t)| \leqslant \frac{1}{\pi} \int_{-T}^{T} \left| \frac{\psi(\xi) - e^{-\xi^2/2}}{\xi} \right| d\xi + \frac{4}{T}. \tag{9.1}$$

First note that we may assume $\eta < 1$, otherwise the theorem is trivial. Let I be the integral in (9.1) and set $a := \eta^{-1/3}$. We bound the I by breaking the integral into two ranges: $|\xi| \in [0,a]$ and $|\xi| \in [a,T]$. For $|\xi| \leqslant a$, we bound the integrand

$$\left| \frac{\psi(\xi) - e^{-\xi^2/2}}{\xi} \right| = e^{-\xi^2/2} \left| \frac{e^{R(\xi)} - 1}{\xi} \right| \leqslant 4\eta e^{-\xi^2/2} |\xi|^2, \tag{9.2}$$

since $|e^z - 1| \leqslant 4|z|$ for $|z| \leqslant 1$. For $a \leqslant |\xi| \leqslant T$, we use the fact that $a = \eta^{-1/3} \geqslant 1$ along with the triangle inequality to bound the integrand

$$\left| \frac{\psi(\xi) - e^{-\xi^2/2}}{\xi} \right| \leqslant e^{-\xi^2/4} |e^{-\xi^2/4} - e^{-\xi^2/4 + R(\xi)}| \leqslant 2e^{-\xi^2/4} \cdot e^{-\xi^2/4 + |R(\xi)|},$$

where

$$-\frac{\xi^2}{4} + |R(\xi)| = |\xi|^2 \Big(-\frac{1}{4} + \eta|\xi| \Big) \leqslant -\frac{\xi^2}{8}$$

and the last inequality holds due to the fact that $\eta|\xi| \leqslant \eta T \leqslant 1/8$, by the choice of T. So, for $|\xi| \in [a,T]$, we have

$$\left| \frac{\psi(\xi) - e^{-\xi^2/2}}{\xi} \right| \leqslant 2e^{-\xi^2/4} e^{-\xi^2/8} \leqslant 2e^{-\xi^2/4} e^{-a^2/8} \leqslant 32\eta e^{-\xi^2/4} \tag{9.3}$$

due to the fact that $\exp(-|x|^{2/3}/8) \leqslant 16/|x|$. Using (9.2) and (9.3), we can bound I by

$$I \leqslant 8\eta \int_0^a e^{-\xi^2/2} |\xi|^2 \, d\xi + 64\eta \int_a^T e^{-\xi^2/4} \, d\xi \leqslant 2^9 \eta,$$

where we have used the facts

$$\int_{-\infty}^{\infty} e^{-t^2/2} |t|^2 \, dt = \sqrt{2\pi} \quad \text{and} \quad \int_{-\infty}^{\infty} e^{-t^2/4} \, dt = 2\sqrt{2\pi}.$$

Putting this together with (9.1) gives the bound

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X \leqslant t) - \mathbb{P}(Z \leqslant t)| \leqslant \pi^{-1} I + 32 \max\{\eta, \tau^{-1}\} \leqslant 2^9 \max\{\eta, \tau^{-1}\},$$

as desired. ■

We note that Lemma 9.1, along with our main technical Lemma 8.1, implies the following general result.

Corollary 9.3. *For $\varepsilon \in (0,1)$ and $b \geqslant 0$, let X be a random variable with logarithmic potential u_X and normalized cumulant sequence $(a_j)_{j\geqslant 2}$. If u_X is b-decreasing and harmonic in $B(1,\varepsilon)$ and*

$$\sum_{j\geqslant 2} |a_j| \Big(\frac{\varepsilon}{32}\Big)^j > b, \tag{9.4}$$

then[13]

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O\Big(\frac{1}{\varepsilon\sigma}\Big),$$

where $Z \sim N(0,1)$ is a standard normal variable.

9.2. *Proof of Theorem 1.4*

We are now ready to prove our main theorem on random variables with roots avoiding a sector.

Proof of Theorem 1.4. Let $X \in \{0,\dots,n\}$ be a random variable for which its probability generating function f_X has no roots in the sector $S(\delta)$. This means that its logarithmic potential $u(z) = u_X(z)$ is a weakly positive, symmetric and harmonic function on $S(\delta)$. Also, since f_X is a polynomial, we have that

$$u(z) = O(\log|z|) \quad \text{and} \quad u\Big(\frac{1}{z}\Big) = O(\log|z|)$$

as $z \to \infty$. Finally, note that we may assume $\sigma > 0$, otherwise the statement of Theorem 1.4 is meaningless.

We first look to apply Lemma 4.1 to show that u is decreasing in a neighborhood of $1 \in \mathbb{C}$; that is, b-decreasing for $b = 0$. For this, note that for all $R > 1$ we have that $S_R(\delta/2) \subseteq S(\delta/2)$ and so u is harmonic *in a neighborhood* of $S_R(\delta/2)$. Set $r := 2$ and we check, in accordance with (4.1), that

$$\Big(\frac{2}{R}\Big)^{1/\delta} \max_{z\in S_R^*(0,\delta)} |u(z)| = O(R^{-1/\delta}\log R) \to 0,$$

as $R \to \infty$, due to the growth condition on u. Thus, we may apply Lemma 4.1 to learn that u is b-decreasing in $S_2(0,\delta/2)$ for every $b \geqslant 0$, and therefore is *decreasing*.

We now look to apply Corollary 9.3 to finish the proof of the theorem. For this, we only have to check the condition at (9.4), which easily follows from the fact that $\sigma > 0$. Since u is 0-decreasing and harmonic in $B(1,\delta/4)$, we may apply Corollary 9.3 with $\varepsilon = \delta/4$ to finish the proof. ■

[13] Here, the implicit constant Corollary 9.3 can be taken to be 2^{3255}.

9.3. *Proof of Theorem* 1.2

The proof of Theorem 1.2 is similar to Theorem 1.4. In the proof of Theorem 1.2, we work in a tiny truncated sector which we place in the zero-free ball $B(1,\varepsilon)$ and then estimate $u_X(z)$ on the ends $S_R^*(\delta)$ of the boundary $\partial S_R(\delta)$. This estimate is the content of the following lemma.

Lemma 9.4. *For $\delta \in (0,1/2)$, let $X \in \{0,\dots,n\}$ be a random variable for which f_X has no roots in the ball $B(1,\delta)$. For $\varepsilon \in (0,\delta/4)$ and $R = 1+\delta/4$, we have*

$$\max_{z\in S_R^*(\varepsilon)} |u_X(z)| \leqslant 7n\log\left(\frac{4}{\delta}\right).$$

Proof. Note that $z \in S_R^*(\varepsilon)$ implies $|z-1| \leqslant \delta/2$. Write $z = 1+s$ and expand

$$u_X(1+s) = \log|f_X(1+s)| = \log\left|\prod_\lambda\left(1+\frac{s}{1-\lambda}\right)\right| = \sum_\lambda \log\left|1+\frac{s}{1-\lambda}\right|.$$

Since $|1-\lambda| \geqslant \delta$ and $|s| \leqslant \delta/2$, the triangle inequality implies

$$\max_{z\in S_R^*(\varepsilon)} |u_X(z)| \leqslant n \max_{|y|\leqslant 1/2} |\log|1+y|| \leqslant 2n. \qquad \blacksquare$$

Proof of Theorem 1.2. Let $X \in \{0,\dots,n\}$ be a random variable for which its probability generating function f_X has no zeros in the ball $B = B(1,\delta)$, for $\delta \in (0,1/2)$. Note that we may assume that $\sigma > 2^{10}\delta^{-1}\log n$ otherwise (1.1) is trivial, with implicit constant 2^{10}.

Using the fact that there are no roots in B implies that $u(z) = u_X(z)$ is harmonic in the ball B. We now work in a thin, truncated sector inside of B. In particular, we set $\varepsilon = \delta/(32\log n)$, $R = 1+\delta/4$, $r = 1+\varepsilon$ and work in the sector $S := S_R(\varepsilon)$. Since $S_R(\varepsilon) \subseteq B$, we have that u is harmonic in a neighborhood of $S_R(\varepsilon)$.

Now choose $b = 1$ and, looking to apply Lemma 4.1, we verify (4.1). First note that for $n \geqslant 3$,

$$\frac{1}{\varepsilon}\log\frac{R}{r} \geqslant (32\log n)\delta^{-1}(\log R - \log r) \geqslant \log n\left(\frac{\delta}{4}-\frac{\delta}{8}\right) \geqslant 4\log n,$$

and therefore

$$\exp\left(-\frac{1}{\varepsilon}\frac{R}{r}\right)\max_{z\in S_R^*(0,\varepsilon)}|u(z)| \leqslant \frac{1}{n^4}(2n) \leqslant \frac{2}{n^3}, \tag{9.5}$$

using Lemma 9.4. Thus, (9.5) is at most $1\cdot(3/8)$, for $n \geqslant 3$, and thus we may apply Lemma 4.1 to conclude that u is 1-decreasing in $B(1,\varepsilon)$.

We now look to apply Corollary 9.3. Since u is 1-decreasing, weakly positive and harmonic in $B(1,\varepsilon)$, we only need to check condition (9.4) for $b = 1$. This is easily done as

$$\sum_{j\geqslant 2}|a_j|\left(\frac{\varepsilon}{32}\right)^j \geqslant \left(\frac{\varepsilon\sigma}{32}\right)^2 \geqslant \left(\frac{\sigma\delta}{2^{10}\log n}\right)^2 > 1 = b,$$

where we have used the assumption that $\sigma > 2^{10}\delta^{-1}\log n$, since the statement is trivial otherwise. We now apply Corollary 9.3 with $\varepsilon = \delta/(32\log n)$ to complete the proof. $\blacksquare$

10. Multivariate central limit theorem for strong Rayleigh measures

In this section, we prove that strong Rayleigh distributions satisfy central limit theorems. For a $d \times d$ positive semi-definite matrix A and a vector $\mu \in \mathbb{R}^d$, we define $N(\mu, A)$ to be the multivariate Gaussian random variable with mean μ and covariance matrix A. To prove that a random variable $X \in \mathbb{R}^d$ is a multivariate normal distribution, we show that "many" of its one-dimensional projections $\langle X, v \rangle$ are Gaussian. We will then apply a variant of the famous Cramér–Wold theorem which will allow us to conclude that X itself must be a *multivariate* Gaussian.

The key connection between stable polynomials and univariate polynomials that have no roots in a sector comes from the following fundamental observation, first made by Ghosh, Liggett and Pemantle [28].

Lemma 10.1. *Let $Y \in \mathbb{Z}^d$ be a finitely supported random variable with real-stable probability generating function f_Y. If $v = (v_1, \dots, v_d) \in \mathbb{Z}^d_{\geqslant 0}$, then the probability generating function of $\langle v, Y \rangle$ has no zeros in the sector $S(\pi/\|v\|_\infty)$.*

Proof. Let f_Y be the probability generating function of Y, let f_v be the probability generating function of $Y(v) := \langle v, Y \rangle$, and let ζ be a root of f_v. First note that

$$f_v(z) = f_Y(z^{v_1}, \dots, z^{v_d}). \tag{10.1}$$

Now write $\zeta = re^{i\theta}$ for $r > 0$ and $\theta \in [-\pi, \pi]$. Since $f_v \in \mathbb{R}[z]$, we may additionally assume that $\theta \in [0, \pi]$, by possibly replacing $re^{i\theta}$ by its conjugate. From (10.1), we see that $(r^{v_1} e^{iv_1\theta}, \dots, r^{v_d} e^{iv_d\theta})$ is a root of $f_Y(z_1, \dots, z_d)$ and therefore, by the real-stability of f, there is some $i \in [d]$ with $\sin(v_i\theta) < 0$ and therefore $\theta \in (\pi/v_i, \pi] \subseteq (\pi/\|v\|_\infty, \pi]$. ∎

Lemma 10.1 allows us to use Theorem 1.4 to show that all the projections of a strong Rayleigh distribution are approximately normal, in non-negative integer directions. Note that we will take the degenerate normal $N(0, 0)$ to also be a normal random variable: indeed, this measure is simply the point mass at 0.

Lemma 10.2. *For each $n \geqslant 1$, let $X_n \in \{0, \dots, n\}^d$ be a strong Rayleigh distribution with mean μ_n, covariance matrix A_n and maximum variance σ_n^2. Put $X_n^* = (X_n - \mu_n)\sigma_n^{-1}$. If $\sigma_n \to \infty$ and $\sigma_n^{-2} A_n \to A$, then for all $v \in \mathbb{Z}^d_{\geqslant 0}$, we have that*

$$\langle X_n^*, v \rangle \to N(0, v^\mathsf{T} A v),$$

in distribution.

Proof. Let us put $Y_n(v) := \langle X_n, v \rangle = v_1 X_1 + \cdots + v_d X_d$. Note that $\mathbb{E} Y_n(v) = \langle v, \mu_n \rangle$ and that

$$\operatorname{Var}(Y_n(v)) = v^\mathsf{T} A_n v.$$

From Lemma 10.1, we see that the probability generating function $f_{Y_n(v)}$ of $Y_n(v)$ has no roots in the sector $S(\delta)$, where $\delta = \pi/\|v\|_\infty$.

There are two cases: when v is in the null space of A and when v is not in the null space of A. Let us first assume that $Av \neq 0$. In this case, we have

$$\lim_n \mathrm{Var}(Y_n(v))\sigma_n^{-2} = \lim_n v^\top(\sigma_n^{-2}A_n)v = v^\top Av \neq 0,$$

and in particular $\mathrm{Var}(Y_n(v)) \to \infty$. Thus, we may apply our central limit theorem for random variables avoiding a sector, Corollary 1.5, to see that

$$\frac{Y_n(v) - \mathbb{E}Y_n(v)}{(v^\top A_n v)^{1/2}} \to N(0,1)$$

and therefore

$$\frac{\sigma_n}{(v^\top A_n v)^{1/2}}\langle X_n^*, v\rangle = \frac{Y_n(v) - \mathbb{E}Y_n(v)}{(v^\top A_n v)^{1/2}} \to N(0,1).$$

Since $(v^\top A_n v)^{1/2}\sigma_n^{-1}$ tends to a constant as $n \to \infty$, it follows that

$$\langle X_n^*, v\rangle \to N(0, v^\top Av),$$

as desired.

In the other case, we have that $Av = 0$ and therefore $\mathrm{Var}(Y_n(v))\sigma_n^{-2} \to 0$. So, for all $x > 0$, we may apply Chebyshev's inequality to see that

$$\mathbb{P}(|Y_n(v) - \mathbb{E}Y_n(v)| \geqslant x\sigma_n) \leqslant \mathrm{Var}(Y_n(v))(x\sigma_n)^{-2} = o(1).$$

This simply means that $(Y_n(v) - \mathbb{E}Y_n(v))\sigma_n^{-1}$ tends to a point mass at zero, in distribution. So trivially,

$$\langle X_n^*, v\rangle \to N(0,0) = N(0, v^\top Av),$$

as desired. ■

To "lift" this information about the projected random variables, we appeal to a theorem of Cuesta-Albertos, Fraiman and Ransford [20], which will allow us to conclude that the distribution of our multivariate random variable is approximately normal from the fact that "many" of its projections are normal. To properly state this result, we let ν be a Borel probability measure on $\mathbb{R}^d$ and, for $v \in \mathbb{R}^d$, we define the measure ν_v to be the "projected" measure on $\mathbb{R}$ by

$$\nu_v(B) = \nu(\{x \in \mathbb{R}^d : \langle v, x\rangle \in B\})$$

for every Borel set $B \subseteq \mathbb{R}$. If $\widetilde{\nu}$ is another Borel probability measure on $\mathbb{R}^d$, we define $\Pi(\nu, \widetilde{\nu}) \subseteq \mathbb{R}^d$ to be the set of $v \in \mathbb{R}^d$ for which $\widetilde{\nu}_v = \nu_v$. In this notation, the classical Cramér–Wold theorem [19] says that if $\nu, \widetilde{\nu}$ are Borel probability measures such that $\Pi(\nu, \widetilde{\nu}) = \mathbb{R}^d$, then $\nu = \widetilde{\nu}$. Cuesta-Albertos, Fraiman and Ransford [20] have sharpened this result by showing that it is enough for $\Pi(\nu, \widetilde{\nu})$ to not be contained in the zero-set of a polynomial. We shall only make use of the following corollary of this theorem.

Corollary 10.3. *For $d \geqslant 1$, let A be $d \times d$ positive semi-definite matrix, and let ν_A be the Gaussian distribution on $\mathbb{R}^d$ with covariance matrix A and mean zero. If ν is a measure for which $\Pi(\nu_A, \nu) \supseteq \mathbb{Z}^d_{\geqslant 0}$, then $\nu_A = \nu$.*

Recall that a sequence of Borel probability measures ν_n on $\mathbb{R}^d$ is said to be *tight*, if for every $\varepsilon > 0$ there exists $R = R(\varepsilon) > 0$ such that the ball $B(0, R)$ satisfies $\nu_n(B(0, R)) > 1 - \varepsilon$, for all sufficiently large n. For the proof of Theorem 1.6, we need two basic facts about tight sequences of measures (see, e.g., [9, Theorem 25.10]). For each $n \geqslant 1$, let $X_n \in \mathbb{R}^d$ be a random variable, with finite mean μ_n, covariance matrix A_n and maximum variance $\sigma_n^2 \in (0, \infty)$. If ν_n is the law of $X_n^* := (X_n - \mu_n)\sigma_n^{-1}$, then ν_n is a tight sequence of measures. We also need that if ν_n is a tight sequence of Borel probability measures on $\mathbb{R}^d$, then there exist a subset $S \subseteq \mathbb{N}$ and a Borel probability measure ν for which $\nu_n \to \nu$, weakly, for $n \in S$.

We can now finish the proof of Theorem 1.6.

Proof of Theorem 1.6. For each $n \geqslant 1$, let $X_n \in \{0, \ldots, n\}^d$ be a random variable with mean μ_n, covariance matrix A_n, maximum variance σ_n^2, and let ν_n be the law of $X_n^* := (X_n - \mu_n)\sigma_n^{-1}$. We have that $\sigma_n^{-2} A_n \to A$, for some (non-zero) matrix A.

Let ν_A denote the law of $N(0, A)$. We show that every subsequence has a further subsequence that converges to ν_A, which is enough to conclude that $\nu_n \to \nu_A$. For this, let $S \subset \mathbb{N}$; by tightness of $\{\nu_n\}$, we may find $S' \subset S$ so that along S' we have $\nu_n \to \nu'$ for some measure ν'. Convergence in distribution together with Lemma 10.2 imply that $\Pi(\nu_A, \nu') \supseteq \mathbb{Z}^d_{\geqslant 0}$. Corollary 10.3 then implies $\nu' = \nu_A$. This completes the proof. ∎

Remark 10.4. We point out that our results here are easily generalized beyond real-stable polynomials to other situations where f_X satisfies a certain "half-plane property". To take a well-known example, we say that a polynomial is *Hurwitz stable* if it has no roots in

$$\{(z_1, \ldots, z_d) \in \mathbb{C}^d : \mathrm{Re}(z_i) > 0 \text{ for all } i\}.$$

Our work here implies a version of Theorem 1.6 in the case that f_X is Hurwitz stable. In fact, the only point to check is Lemma 10.1 and the rest of the proof proceeds in the same way.

More generally, let $\phi \in (0, 2\pi)$ and define

$$H_\phi := \{(z_1, \ldots, z_d) \in \mathbb{C}^d : \arg(z_i) \in [0, \phi] \text{ for all } i\}.$$

We say that a polynomial f is H_ϕ*-stable* if it has no roots in H_ϕ. It is not hard to see that our results imply a central limit theorem for a sequence of random variables X_n when the f_{X_n} are H_ϕ-stable polynomials and $\sigma_n \to \infty$.

11. Sharpness of results

In this section, we show that our quantitative results (Theorems 1.2 and 1.4) are sharp up to the implied constants. From this it will also follow that the conditions in the limit theorems, Corollaries 1.3 and 1.5, are best possible.

Our constructions follow from a few simple observations. This first observation gives a cheap bound on the discrepancy between a discrete random variable and the standard normal distribution. Recall that we use the notation X^* to denote $(X-\mu)\sigma^{-1}$ for a random variable X.

Observation 11.1. *Let $X \in \mathbb{Z}$ be a random variable with mean $\mu < \infty$ and standard deviation*[14] $\sigma \in [2^{-3}, \infty)$. *Then*

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{e^{-16}}{\sigma},$$

where $Z \sim N(0,1)$.

Proof. Note that since $X \in \mathbb{Z}$, we have

$$X^* := (X-\mu)\sigma^{-1} \in \frac{1}{\sigma}(\mathbb{Z}-\mu).$$

Find values $a, b \in \mathbb{Z} - \mu$ such that $b - a = 1$, $a \leqslant 0$ and $b \geqslant 0$. Then

$$\mathbb{P}\Big(X^* \in \frac{1}{\sigma}(a,b)\Big) = 0$$

while

$$\mathbb{P}\Big(Z \in \frac{1}{\sigma}(a,b)\Big) = \frac{1}{(2\pi)^{1/2}}\int_{a/\sigma}^{b/\sigma} e^{-s^2/2}\mathrm{d}s \geqslant \frac{1}{(2\pi)^{1/2}}\int_0^{1/2\sigma} e^{-s^2/2}\mathrm{d}s \geqslant \frac{e^{-8}}{2\sqrt{2\pi}\sigma},$$

where $Z \sim N(0,1)$ and we have used that one of $|a|$ or $|b|$ must be at least $1/2$. This allows us to obtain a lower bound on the maximum discrepancy between the two cumulative distribution functions. We have

$$\begin{aligned}&2\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \\ &\qquad \geqslant \Big|\mathbb{P}\Big(X^* \leqslant \frac{b}{\sigma}\Big) - \mathbb{P}\Big(Z \leqslant \frac{b}{\sigma}\Big)\Big| + \Big|\mathbb{P}\Big(X^* \leqslant \frac{a}{\sigma}\Big) - \mathbb{P}\Big(Z \leqslant \frac{a}{\sigma}\Big)\Big|,\end{aligned}$$

which is at least

$$\mathbb{P}\Big(Z \in \frac{1}{\sigma}(a,b)\Big) - \mathbb{P}\Big(X^* \in \frac{1}{\sigma}(a,b)\Big) \geqslant \frac{e^{-8}}{2\sqrt{2\pi}\sigma},$$

by the triangle inequality. Lower bounding the constant by e^{-16} completes the proof. ■

The next basic observation records a key "trick" in our constructions. It says that the transformation $X \mapsto k \cdot X$ does not change the maximum discrepancy with a normal. However, the standard deviation *increases* as $\sigma(k\cdot X) = k\sigma(X)$.

[14] The 2^{-3} is an arbitrary choice, we just needed a sufficiently small number for the application of Theorem 11.3.

Observation 11.2. *Let $Y \in \mathbb{Z}$ be a random variable with finite mean μ and standard deviation $\sigma \in [2^{-3}, \infty)$. For $k > 0$, let $X = k \cdot Y$. Then*

$$\sup_{t \in R} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{ck}{\sigma(X)},$$

where we can take $c = e^{-16}$.

Proof. Note that $\sigma(X) = k\sigma(Y)$ and $\mathbb{E}X = k\mathbb{E}Y$, thus

$$X^* = (X - \mathbb{E}\, X)\sigma(X)^{-1} = (Y - \mathbb{E}Y)\sigma(Y)^{-1} = Y^*,$$

and so

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = \sup_{t \in \mathbb{R}} |\mathbb{P}(Y^* \leqslant t) - \mathbb{P}(Z \leqslant t)|.$$

Thus, applying Observation 11.1 to Y yields

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{e^{-16}}{\sigma(Y)} = \frac{ck}{\sigma(X)},$$

as desired. ∎

Our constructions for both theorems in this section are achieved simply by applying Observation 11.2 to an appropriate "seed" random variable. For Theorem 11.3, we make use of a simple class of random variables. If $\theta \in [\pi/2, \pi]$, the polynomial

$$P_{\rho,\theta}(z) = (z - \rho e^{i\theta})(z - \rho e^{-i\theta}) = z^2 - 2\rho(\cos\theta)z + \rho^2$$

has non-negative coefficients and therefore $P_{\rho,\theta}(z)(P_{\rho,\theta}(1))^{-1}$ is the probability generating function of a random variable, which we shall denote by $Y_{\rho,\theta}$.

Now note that for each fixed θ, as $\rho \geqslant 1$ increases, $\mathrm{Var}(Y_{\rho,\theta})$ decreases as a continuous function of ρ. Further, each random variable is non-degenerate for $\rho \in [1, \infty)$, implying $\mathrm{Var}(Y_{\rho,\theta}) > 0$. Since we also have $\lim_{\rho \to \infty} \mathrm{Var}(Y_{\rho,\theta}) = 0$, there exists some $a(\theta) > 0$ such that $\{\mathrm{Var}(Y_{\rho,\theta})\}_{\rho \geqslant 1} \supseteq [0, a(\theta)]$.

Theorem 11.3. *For every $\delta \in (0, \pi]$ and $\sigma > 0$ with $\delta\sigma \geqslant 1$, there exists a random variable $X \in \mathbb{Z}_{\geqslant 0}$, which is supported on finitely many integers, with standard deviation σ and probability generating function f_X for which $\delta = \min_{\{\zeta : f(\zeta) = 0\}} |\arg(\zeta)|$ and*

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{c}{\delta\sigma},$$

where we can take $c = e^{-16}$.

Proof. Let (σ, δ) be given. As $\bigcup_{j \geqslant 0} [\pi/2^{j+1}, \pi/2^j] = (0, \pi]$, we may write $\delta = \theta/k$ for some $k \in \mathbb{N}$ and $\theta \in [\pi/2, \pi]$ and note that $1 \leqslant \delta\sigma = (\theta\sigma)/k$. We start by constructing a random variable Y with standard deviation $\sigma/k \geqslant 2^{-3}$ and $\min_\zeta |\arg(\zeta)| = \theta$. We then finish by applying Observation 11.2.

For ρ, m to be chosen later, let Y_i be independent copies of $Y_{\rho,\theta}$ and let

$$Y = \sum_{i=1}^{m} Y_i.$$

Of course, $\sigma(Y) = m^{1/2}\sigma(Y_{\rho,\theta})$ and therefore, from the discussion that precedes Theorem 11.3, we may choose m, ρ such that $m^{1/2}\sigma(Y_{\rho,\theta}) = \sigma/k$. Moreover, every root ζ of the probability generating function $f_Y = (f_{Y_{\rho,\theta}})^m$ of Y has $\arg(\zeta) \in \{-\theta, \theta\}$.

Finally, set $X = k \cdot Y$. The probability generating function of X is $f_X(z) = f_Y(z^k)$, and thus the roots ζ of f_X satisfy

$$\arg(\zeta) \in \left\{\pm\frac{\theta}{k} + \frac{2\pi\ell}{k} \bmod 2\pi : \ell \in \{0, \ldots, k-1\}\right\},$$

and therefore $\min_{\{\zeta : f_X(\zeta)=0\}} |\arg(\zeta)| = \theta/k = \delta$. From Observation 11.2, we have that

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{e^{-16}k}{\sigma(X)} \geqslant \frac{e^{-16}}{\delta\sigma(X)},$$

where the last inequality follows from the fact that $k\delta = \theta \in [\pi/2, \pi]$. This completes the proof. ∎

The following shows that Theorem 1.2 is sharp. Here, we apply Observation 11.2 to a sum of Bernoulli random variables.

Theorem 11.4. *For $n \geqslant 1$, $\delta > 0$ and $\sigma^2 \in [1, n^{0.9})$ satisfying $\log n/(\delta\sigma) \leqslant 1$, there exists a random variable $X \in \{0, \ldots, n\}$ with standard deviation σ such that*

$$\min_{\{\zeta : f_X(\zeta)=0\}} |1-\zeta| \geqslant \delta \quad \textit{and} \quad \sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{c\log n}{\delta\sigma}.$$

Proof. Let $\{Y_i\}_{i\geqslant 1}$ be independent and identically distributed Bernoulli random variables, where $p := \mathbb{P}(Y_i = 1)$ will be chosen later. Of course, we have that $\mathrm{Var}(Y_i) = p(1-p)$ and the probability generating function of Y_i is $pz + (1-p)$. We set

$$Y = \sum_{i=1}^{\lfloor n/k \rfloor} Y_i,$$

with $k := \lfloor \log n/(100\delta) \rfloor$ and note that $\mathrm{Var}(Y) = \lfloor n/k \rfloor p(1-p)$. We define $X := kY$ and set $p = n^{-\alpha}$ with $\alpha \in [0.01, 1)$ to be chosen later. To apply Observation 11.2 to X, we require that $\mathrm{Var}(Y) \geqslant 1/8$; and so we impose the condition

$$n^{1-\alpha} \geqslant \frac{\log n}{100\delta}$$

to guarantee this. Now,

$$\mathrm{Var}(X) = k^2 \left\lfloor \frac{n}{k} \right\rfloor p(1-p)$$

is a continuous function of α as $p = n^{-\alpha}$. Also

$$\mathrm{Var}(X) \leqslant knp(1-p) \leqslant \frac{\log n}{100\delta} n^{1-\alpha}(1-n^{-\alpha})$$

and

$$\mathrm{Var}(X) \geqslant knp(1-p) - k^2 p(1-p) \geqslant \frac{\log n}{200\delta} n^{1-\alpha}(1-n^{-\alpha}),$$

where the last line holds when $n \geqslant \log n/(200\delta) \geqslant k/2$, which always holds for us since $n \geqslant \sigma \geqslant \log n/\delta$, by hypothesis. So as $\alpha \in [0.01, 1)$ varies subject to $n^{1-\alpha} \geqslant \log n/(100\delta)$, $\mathrm{Var}(X)$ ranges over a set containing the interval $[\delta^{-2}(\log n)^2, n^{0.9}]$ and as $\log n/(\delta\sigma) \leqslant 1$ and $\sigma^2 < n^{0.9}$, we may select $\alpha \in [0.01, 1)$ so that $\mathrm{Var}(X) = \sigma^2$.

Now note that $\deg(f_X) = k\lfloor n/k \rfloor \leqslant n$ and thus $X \in \{0, \dots, n\}$. Since $f_X(z) = f_Y(z^k)$, the roots ζ of f_X are of the form $\zeta = \beta((1-p)/p)^{1/k}$, where $|\beta| = 1$, which allows us to bound

$$\min_{\zeta} |1-\zeta| \geqslant |1 - e^{\alpha \log n/k}| = |1 - e^{\alpha \log n/\lfloor \log n/(100\delta)\rfloor}| \geqslant \delta.$$

Applying Observation 11.2, we see that

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{ck}{\sigma(X)} \geqslant \frac{c\log n}{100\delta\sigma} \geqslant \frac{C\log n}{\delta\sigma},$$

as desired. ■

12. General distributions

In this brief section, we discuss how to apply our results to random variables that take values in $\mathbb{R}$, rather than just in $\{0, \dots, n\}$. In short, everything for Theorem 1.4 extends rather naturally, but a few extra complications arise.

The first task is to fix an appropriate notion of the probability generating function of X. Luckily, there is already a standard definition in this situation. First set $z^r := \exp(r \log z)$, for all $r \in \mathbb{R}$, where "log" denotes the standard branch of the logarithm; then define

$$f_X(z) := \mathbb{E}_X z^X,$$

for all $z \in \mathbb{C} \setminus \mathbb{R}_{\leqslant 0}$, to be the probability generating function of X.

We now happen upon a feature of the more general set-up: f_X does not necessarily *exist* for all z and therefore it may not make any sense to discuss the zeros of $f_X(z)$ at all. To ensure the existence of f_X, for all $z \notin \mathbb{R}_{\leqslant 0}$, it is enough to impose the condition $f_Z(\rho) < \infty$ for all $\rho > 0$. With this assumption in hand, Morera's theorem shows that f is analytic as well.

Lemma 12.1. *Let $X \in \mathbb{R}$ be a random variable and let f_X be its probability generating function. If $f(\rho) < \infty$ for all $\rho > 0$, then $f_X(z)$ is analytic in $\mathbb{C} \setminus \mathbb{R}_{\leqslant 0}$.*

Proof. Let γ be a piecewise C^1 closed contour in $\mathbb{C} \setminus \mathbb{R}_{\leqslant 0}$. Since $\log z$ is analytic in $\mathbb{C} \setminus \mathbb{R}_{\geqslant 0}$, Fubini's theorem shows

$$\oint_\gamma f_X(z)\,\mathrm{d}z = \mathbb{E}_X \oint_\gamma \exp(X \log z)\,\mathrm{d}z = 0.$$

Morera's theorem then implies f_X is analytic in $\mathbb{C} \setminus \mathbb{R}_{\geqslant 0}$. ■

A second subtlety concerns the asymptotic growth of the logarithmic potential

$$u_X(z) := \log|f(z)|$$

for $|z|$ very large and very small. For $\kappa, \delta > 0$, we say that u satisfies the (κ, δ)-*growth condition* if we have

$$\lim_{|z|\to\infty} \frac{|u_X(z)|}{|z|^\kappa} = 0 \quad \text{and} \quad \lim_{|z|\to\infty} \frac{|u_X(1/z)|}{|z|^\kappa} = 0, \tag{12.1}$$

where the limits are taken with $z \in S(\delta)$.

In previous sections, we could ignore (12.1), as u_X trivially satisfies the (κ, δ)-growth condition for all $\kappa > 0$ when f_X is a polynomial. Here, however, we are forced to take the rate of growth into account, as it directly affects the convergence to a normal distribution.

We now state our main general theorem for zero-free sectors of probability generating functions.

Theorem 12.2. *For $\delta > 0$ and $\kappa > 0$, let $X \in \mathbb{R}$ be a random variable with probability generating function f_X for which $f(\rho)$ is defined for all $\rho \in \mathbb{R}_{\geqslant 0}$. If u_X satisfies the (κ, δ)-growth condition and f_X has no zeros in $S(\delta)$, then*[15]

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| = O\Big(\frac{\max\{\delta^{-1}, \kappa\}}{\sigma}\Big),$$

where $Z \sim N(0, 1)$.

Again, this theorem is sharp with respect to the dependence on κ, δ and σ, as we shall see in Section 12.2.

12.1. Proof of Theorem 12.2

We first notice that many of the properties of the logarithmic potential of X easily carry over to this more general setting. Indeed, if f_X is zero-free in the sector $S(\delta)$, then $u(z)$ is harmonic in this sector. Also $u(z)$ is symmetric and weakly positive. We also have that $f_X(1) = 1$ and therefore $u_X(1) = 0$. With these observations at hand, we may prove Theorem 12.2 as we proved Theorem 1.4.

[15] The implicit constant may be taken to be 2^{3258}.

Proof of Theorem 12.2. Let $X \in \mathbb{R}$ be a random variable with probability generating function f_X, satisfying $f_X(\rho) < \infty$ for all $\rho > 0$, that is, zero-free in the sector $S(\delta)$ and such that u_X satisfies the (κ, δ)-growth condition. By the discussion above, we know that the logarithmic potential $u = u_X$ is harmonic, symmetric and weakly positive in $S(\delta)$. Also note that we may assume that $\sigma > 0$, otherwise the statement of the Theorem 12.2 is meaningless.

We now choose $\varepsilon = \min\{\delta/2, 1/(2\kappa)\}$ and note that u is a weakly positive, symmetric and harmonic function on the smaller region $S(\varepsilon)$. Now, looking to apply Lemma 4.1, we set $r := 2$ and note that

$$\left(\frac{2}{R}\right)^{1/\varepsilon} \max_{z \in S_R^*(\varepsilon)} |u(z)| = O(R^{-1/\varepsilon+\kappa}) = O(R^{-\kappa}) \to 0$$

as $R \to \infty$. Thus, we may apply Lemma 4.1 to learn that u is decreasing in $S(\varepsilon/2)$. This implies that u is decreasing in $B(1, \varepsilon/4)$. Since $\sigma > 0$, u satisfies the conditions of Corollary 9.3, which finishes the proof. ∎

12.2. Proof of the sharpness of Theorem 12.2

Theorem 12.3. *Let $\kappa, \delta \in (0, \pi)$ and σ be such that $\sigma \cdot \min\{\delta, 1/\kappa\} \geqslant 1$. Then there exists a random variable $X \in \mathbb{Z}$ with standard deviation σ such that u_X is harmonic in $S(\delta)$, u_X satisfies the (κ, δ)-growth condition, and*

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(X^* \leqslant t) - \mathbb{P}(Z \leqslant t)| \geqslant \frac{c}{\sigma} \cdot \max\{\kappa, \delta^{-1}\}. \tag{12.2}$$

Proof. If $\delta \leqslant 1/\kappa$, we apply Theorem 11.3 to obtain a random variable $X \in \mathbb{Z}_{\geqslant 0}$ which has finite support and satisfies (12.2). Here, f_X is a polynomial so

$$\log|f_X| = O(\log|z|) = O(|z|^\kappa).$$

In the case of $\delta > 1/\kappa$, let Y be the Poisson random variable with mean $4\sigma^2/\kappa^2$. Then $Y \in \mathbb{Z}$ with $\sigma(Y) = 2\sigma/\kappa$. Set $X = (\kappa/2) \cdot Y$ and note $\sigma(X) = \sigma$ and

$$u_X(z) = \frac{4\sigma^2}{\kappa^2}(z^{\kappa/2} - 1),$$

which is harmonic in $S(\delta)$ and satisfies the specified growth conditions. Applying Observation 11.2 completes the proof. ∎

Acknowledgments. The authors would like to thank Robin Pemantle for introducing us to these problems and Robert Morris for comments on a draft. The second author is especially grateful to Robin Pemantle and the Mathematics Department at the University of Pennsylvania for hosting him on numerous occasions while this research was being carried out.

Funding. The first author is supported in part by an NSF CAREER grant DMS-2336788 as well as grants DMS-2137623 and DMS-2246624.

References

[1] Amoroso, F., Mignotte, M.: On the distribution of the roots of polynomials. Ann. Inst. Fourier (Grenoble) **46**, 1275–1291 (1996) Zbl 0867.26009 MR 1427125

[2] Anari, N., Oveis Gharan, S., Rezaei, A.: Monte Carlo Markov chain algorithms for sampling strongly Rayleigh distributions and determinantal point processes. In: 29th Annual Conference on Learning Theory, Proc. Mach. Learn. Res. (PMLR) 49, Columbia University, New York, 103–115 (2016)

[3] Anari, N., Oveis Gharan, S., Saberi, A., Srivastava, N.: Approximating the largest root and applications to interlacing families. In: Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 1015–1028 (2018) Zbl 1434.65064 MR 3775854

[4] Axler, S., Bourdon, P., Ramey, W.: Harmonic function theory. 2nd ed., Grad. Texts in Math. 137, Springer, New York (2001) Zbl 0959.31001 MR 1805196

[5] Bender, E. A.: Central and local limit theorems applied to asymptotic enumeration. J. Combin. Theory Ser. A **15**, 91–111 (1973) Zbl 0242.05006 MR 375433

[6] Bender, E. A., Richmond, L. B.: Central and local limit theorems applied to asymptotic enumeration II: Multivariate generating functions. J. Combin. Theory Ser. A **34**, 255–265 (1983) Zbl 0511.05003 MR 700034

[7] Bergweiler, W., Eremenko, A.: Distribution of zeros of polynomials with positive coefficients. Ann. Acad. Sci. Fenn. Math. **40**, 375–383 (2015) Zbl 1332.30012 MR 3310090

[8] Bergweiler, W., Eremenko, A., Sokal, A.: Roots of polynomials with positive coefficients. https://www.math.purdue.edu/~eremenko/dvi/bes.pdf, visited on 16 October 2024 (2013)

[9] Billingsley, P.: Probability and measure. 3rd ed., Wiley Ser. Probab. Math. Statist., John Wiley & Sons, New York (1995) Zbl 0822.60002 MR 1324786

[10] Bilu, Y.: Limit distribution of small points on algebraic tori. Duke Math. J. **89**, 465–476 (1997) Zbl 0918.11035 MR 1470340

[11] Bloch, A., Pólya, G.: On the roots of certain algebraic equations. Proc. Lond. Math. Soc. (2) **33**, 102–114 (1931) Zbl 57.0128.03 MR 1576817

[12] Borcea, J., Brändén, P.: The Lee–Yang and Pólya–Schur programs. I. Linear operators preserving stability. Invent. Math. **177**, 541–569 (2009) Zbl 1175.47032 MR 2534100

[13] Borcea, J., Brändén, P.: The Lee–Yang and Pólya–Schur programs. II. Theory of stable polynomials and applications. Comm. Pure Appl. Math. **62**, 1595–1631 (2009) Zbl 1177.47041 MR 2569072

[14] Borcea, J., Brändén, P., Liggett, T. M.: Negative dependence and the geometry of polynomials. J. Amer. Math. Soc. **22**, 521–567 (2009) Zbl 1206.62096 MR 2476782

[15] Brändén, P., Jonasson, J.: Negative dependence in sampling. Scand. J. Stat. **39**, 830–838 (2012) Zbl 1319.62018 MR 3000852

[16] Cauchy, A. L. B.: Exercises de mathématiques. De Bure frères (1895)

[17] Choe, Y.-B., Oxley, J. G., Sokal, A. D., Wagner, D. G.: Homogeneous multivariate polynomials with the half-plane property. Adv. in Appl. Math. **32**, 88–187 (2004) Zbl 1054.05024 MR 2037144

[18] Conway, J. B.: Functions of one complex variable I. 2nd ed., Grad. Texts in Math. 11, Springer, New York (1978) Zbl 0277.30001 MR 503901

[19] Cramér, H., Wold, H.: Some theorems on distribution functions. J. Lond. Math. Soc. **11**, 290–294 (1936) Zbl 0015.16801 MR 1574927

[20] Cuesta-Albertos, J. A., Fraiman, R., Ransford, T.: A sharp form of the Cramér–Wold theorem. J. Theoret. Probab. **20**, 201–209 (2007) Zbl 1124.60004 MR 2324526

[21] De Angelis, V.: Asymptotic expansions and positivity of coefficients for large powers of analytic functions. Int. J. Math. Math. Sci. 1003–1025 (2003) Zbl 1033.41016 MR 1976089

[22] Dobrushin, R. L.: Estimates of semi-invariants for the Ising model at low temperatures. In: Topics in statistical and theoretical physics, Amer. Math. Soc. Transl. Ser. 2 177, American Mathematical Society, Providence, RI, 59–81 (1996) Zbl 0873.60074 MR 1409170

[23] Erdős, P.: On the uniform distribution of the roots of certain polynomials. Ann. of Math. (2) **43**, 59–64 (1942) Zbl 0060.05502 MR 5947

[24] Erdős, P., Turán, P.: On the distribution of roots of polynomials. Ann. of Math. (2) **51**, 105–119 (1950) Zbl 0036.01501 MR 33372

[25] Feller, W.: An introduction to probability theory and its applications. Vol. II. 2nd ed., John Wiley & Sons, New York (1971) Zbl 0219.60003 MR 270403

[26] Fernández, R., Procacci, A.: Regions without complex zeros for chromatic polynomials on graphs with bounded degree. Combin. Probab. Comput. **17**, 225–238 (2008) Zbl 1171.05017 MR 2396349

[27] Gao, Z., Richmond, L. B.: Central and local limit theorems applied to asymptotic enumeration IV: multivariate generating functions. J. Comput. Appl. Math. **41**, 177–186 (1992) Zbl 0755.05004 MR 1181718

[28] Ghosh, S., Liggett, T. M., Pemantle, R.: Multivariate CLT follows from strong Rayleigh property. In: 2017 Proceedings of the Fourteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO), Society for Industrial and Applied Mathematics, Philadelphia, PA, 139–147 (2017) Zbl 1432.60034 MR 3630942

[29] Godsil, C. D.: Matching behaviour is asymptotically normal. Combinatorica **1**, 369–376 (1981) Zbl 0504.05047 MR 647987

[30] Granville, A.: The distribution of roots of a polynomial. In: Equidistribution in number theory, an introduction, NATO Sci. Ser. II Math. Phys. Chem. 237, Springer, Dordrecht, 93–102 (2007) Zbl 1170.11016 MR 2290496

[31] Gurvits, L.: The van der Waerden conjecture for mixed discriminants. Adv. Math. **200**, 435–454 (2006) Zbl 1093.15011 MR 2200852

[32] Gurvits, L.: Van der Waerden/Schrijver–Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: One theorem for all. Electron. J. Combin. **15**, article no. 66 (2008) Zbl 1182.15008 MR 2411443

[33] Harper, L. H.: Stirling behavior is asymptotically normal. Ann. Math. Statist. **38**, 410–414 (1967) Zbl 0154.43703 MR 211432

[34] Heilmann, O. J., Lieb, E. H.: Theory of monomer-dimer systems. Comm. Math. Phys. **25**, 190–232 (1972) Zbl 0228.05131 MR 297280

[35] Hurwitz, A.: Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt. In: Stability theory (Ascona, 1995), Internat. Ser. Numer. Math. 121, Birkhäuser, Basel, 239–249 (1996) Zbl 0962.01500 MR 1416383

[36] Iagolnitzer, D., Souillard, B.: Lee–Yang theory and normal fluctuations. Phys. Rev. B **19**, 1515–1518 (1979) MR 526177

[37] Jackson, B., Procacci, A., Sokal, A. D.: Complex zero-free regions at large $|q|$ for multivariate Tutte polynomials (alias Potts-model partition functions) with general complex edge weights. J. Combin. Theory Ser. B **103**, 21–45 (2013) Zbl 1257.05068 MR 2995717

[38] Kahn, J.: A normal law for matchings. Combinatorica **20**, 339–391 (2000) Zbl 0963.05111 MR 1774843

[39] Kyng, R., Luh, K., Song, Z.: Four deviations suffice for rank 1 matrices. Adv. Math. **375**, article no. 107366 (2020) Zbl 1460.11106 MR 4137067

[40] Kyng, R., Song, Z.: A matrix Chernoff bound for strongly Rayleigh distributions and spectral sparsifiers from a few random spanning trees. In: 59th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2018, IEEE Computer Society, Los Alamitos, CA, 373–384 (2018) MR 3899605

[41] Lebowitz, J. L., Pittel, B., Ruelle, D., Speer, E. R.: Central limit theorems, Lee–Yang zeros, and graph-counting polynomials. J. Combin. Theory Ser. A **141**, 147–183 (2016) Zbl 1334.05065 MR 3479242

[42] Lee, T. D., Yang, C. N.: Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. Phys. Rev. **87**, 410–419 (1952) Zbl 0048.43401 MR 53029

[43] Li, C., Jegelka, S., Sra, S.: Polynomial time algorithms for dual volume sampling. In: 31st Conference on Neural Information Processing Systems (NIPS 2017), Curran Associates, Inc., 5038–5047 (2017)

[44] Liggett, T. M., Vandenberg-Rodes, A.: Stability on $\{0, 1, 2, \ldots\}^S$: Birth-death chains and particle systems. In: Notions of positivity and the geometry of polynomials, Trends Math., Birkhäuser, Basel, 311–329 (2011) Zbl 1278.60144 MR 3051173

[45] Littlewood, J. E., Offord, A. C.: On the number of real roots of a random algebraic equation. J. Lond. Math. Soc. **13**, 288–295 (1938) Zbl 65.0325.01 MR 1574980

[46] Littlewood, J. E., Offord, A. C.: On the number of real roots of a random algebraic equation. II. Proc. Camb. Philos. Soc. **35**, 133–148 (1939) Zbl 0021.03702

[47] Littlewood, J. E., Offord, A. C.: On the number of real roots of a random algebraic equation. III. Rec. Math. **12(54)**, 277–286 (1943) Zbl 0061.01801 MR 9656

[48] Marcinkiewicz, J.: Sur une propriété de la loi de Gauß. Math. Z. **44**, 612–618 (1939) Zbl 64.0521.02 MR 1545791

[49] Marcus, A. W., Spielman, D. A., Srivastava, N.: Interlacing families I: Bipartite Ramanujan graphs of all degrees. Ann. of Math. (2) **182**, 307–325 (2015) Zbl 1316.05066 MR 3374962

[50] Marcus, A. W., Spielman, D. A., Srivastava, N.: Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem. Ann. of Math. (2) **182**, 327–350 (2015) Zbl 1332.46056 MR 3374963

[51] Michelen, M., Sahasrabudhe, J.: Central limit theorems from the roots of probability generating functions. Adv. Math. **358**, article no. 106840 (2019) Zbl 1472.60046 MR 4021875

[52] Michelen, M., Sahasrabudhe, J.: Anti-concentration of random variables from zero-free regions. Discrete Anal. **2022**, article no. 13 (2022) MR 4503220

[53] Mignotte, M.: Remarque sur une question relative à des fonctions conjuguées. C. R. Acad. Sci. Paris Sér. I Math. **315**, 907–911 (1992) Zbl 0773.31002 MR 1187133

[54] Mörters, P., Peres, Y.: Brownian motion (with an appendix by Oded Schramm and Wendelin Werner). Camb. Ser. Stat. Probab. Math. 30, Cambridge University Press, Cambridge (2010) Zbl 1243.60002 MR 2604525

[55] Oveis Gharan, S., Saberi, A., Singh, M.: A randomized rounding approach to the traveling salesman problem. In: 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science—FOCS 2011, IEEE Computer Society, Los Alamitos, CA, 550–559 (2011) Zbl 1292.68171 MR 2932731

[56] Pemantle, R.: Towards a theory of negative dependence. J. Math. Phys. **41**, 1371–1390 (2000) Zbl 1052.62518 MR 1757964

[57] Pemantle, R.: Hyperbolicity and stable polynomials in combinatorics and probability. In: Current developments in mathematics, 2011, International Press, Somerville, MA, 57–123 (2012) Zbl 1316.62078 MR 3098077

[58] Pemantle, R., Peres, Y.: Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. Combin. Probab. Comput. **23**, 140–160 (2014) Zbl 1326.60032 MR 3197973

[59] Pemantle, R., Wilson, M. C.: Asymptotics of multivariate sequences. II: Multiple points of the singular variety. Combin. Probab. Comput. **13**, 735–761 (2004) Zbl 1065.05010 MR 2095981

[60] Pemantle, R., Wilson, M. C.: Twenty combinatorial examples of asymptotics derived from multivariate generating functions. SIAM Rev. **50**, 199–272 (2008) Zbl 1149.05003 MR 2403050

[61] Peters, H., Regts, G.: On a conjecture of Sokal concerning roots of the independence polynomial. Michigan Math. J. **68**, 33–55 (2019) Zbl 1433.05164 MR 3934603

[62] Regts, G.: Zero-free regions of partition functions with applications to algorithms and graph limits. Combinatorica **38**, 987–1015 (2018) Zbl 1424.05236 MR 3850013

[63] Rota, G.-C., Sharp, D.: Mathematics, philosophy, and artificial intelligence: A dialogue with Gian-Carlo Rota and David Sharp. Los Alamos Science, Spring/Summer (1985)

[64] Ruciński, A.: The behaviour of $\binom{n}{k,\ldots,k,n-ik} c^i / i!$ is asymptotically normal. Discrete Math. **49**, 287–290 (1984) Zbl 0548.05003 MR 743799

[65] Schur, I.: Untersuchungen über algebraische Gleichungen. I: Bemerkungen zu einem Satz von E. Schmidt. Sitzungsber. Preuß. Akad. Wiss. Phys.-Math. Kl. 403–428 (1933) Zbl 59.0911.02

[66] Scott, A. D., Sokal, A. D.: The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. J. Stat. Phys. **118**, 1151–1261 (2005) Zbl 1107.82013 MR 2130890

[67] Scott, A. D., Sokal, A. D.: On dependency graphs and the lattice gas. Combin. Probab. Comput. **15**, 253–279 (2006) Zbl 1138.05323 MR 2195585

[68] Shearer, J. B.: On a problem of Spencer. Combinatorica **5**, 241–245 (1985) Zbl 0587.60012 MR 837067

[69] Soundararajan, K.: Equidistribution of zeros of polynomials. Amer. Math. Monthly **126**, 226–236 (2019) Zbl 1409.42003 MR 3920530

[70] Szegő, G.: Bemerkungen zu einem Satz von E. Schmidt über algebraische Gleichungen. Sitzungsber. Preuß. Akad. Wiss. Phys.-Math. Kl. 86–98 (1934) Zbl 60.0061.01

[71] Wagner, D. G.: Matroid inequalities from electrical network theory. Electron. J. Combin. **11**, article no. 1 (2004) Zbl 1060.05016 MR 2195432

[72] Wagner, D. G.: Weighted enumeration of spanning subgraphs with degree constraints. J. Combin. Theory Ser. B **99**, 347–357 (2009) Zbl 1184.05066 MR 2482953

[73] Yang, C. N., Lee, T. D.: Statistical theory of equations of state and phase transitions. I. Theory of condensation. Phys. Rev. **87**, 404–409 (1952) Zbl 0048.43305 MR 53028

[74] Yogeshwaran, D., Subag, E., Adler, R. J.: Random geometric complexes in the thermodynamic regime. Probab. Theory Related Fields **167**, 107–142 (2017) Zbl 1366.60033 MR 3602843