



# ANNALES DE L'INSTITUT FOURIER

Michael EPSTEIN, BORIS HANIN & ERIK LUNDBERG

**The lemniscate tree of a random polynomial**

Tome 70, n° 4 (2020), p. 1663-1687.

[http://aif.centre-mersenne.org/item/AIF\\_2020\\_\\_70\\_4\\_1663\\_0](http://aif.centre-mersenne.org/item/AIF_2020__70_4_1663_0)

---

© Association des Annales de l'institut Fourier, 2020,

*Certains droits réservés.*



Cet article est mis à disposition selon les termes de la licence  
CREATIVE COMMONS ATTRIBUTION – PAS DE MODIFICATION 3.0 FRANCE.  
<http://creativecommons.org/licenses/by-nd/3.0/fr/>



## THE LEMNISCATE TREE OF A RANDOM POLYNOMIAL

by Michael EPSTEIN, Boris HANIN & Erik LUNDBERG

---

ABSTRACT. — To each generic complex polynomial  $p(z)$  is associated a labeled binary tree (here referred to as a “lemniscate tree”) that encodes the topological type of the graph of  $|p(z)|$ . The branching structure of the lemniscate tree is determined by the configuration (i.e., arrangement in the plane) of the singular components of those level sets  $|p(z)| = t$  passing through a critical point.

In this paper, we ask: how many branches appear in a typical lemniscate tree? We answer this question first for a lemniscate tree sampled uniformly from the combinatorial class of all such trees associated to a generic polynomial of fixed degree and second for the lemniscate tree arising from a random polynomial with i.i.d. zeros. From a more general perspective, these results take a first step toward a probabilistic treatment (within a specialized setting) of Arnold’s program of enumerating algebraic Morse functions.

RÉSUMÉ. — À chaque polynôme complexe générique  $p(z)$  est associé un arbre binaire étiqueté (appelé dans cet article, “arbre de lemniscate”) qui code le type topologique du graphe de  $|p(z)|$ . La structure de ramification de l’arbre de lemniscate est déterminée par la configuration (c’est-à-dire la disposition dans le plan) des composants singuliers de ces ensembles de niveaux  $|p(z)| = t$  passant par un point critique.

Dans cet article, nous nous intéressons à la question suivante: combien de branches apparaissent typiquement dans un arbre de lemniscate ? Nous répondons d’abord à cette question pour un arbre de lemniscate échantillonné uniformément dans la classe combinatoire de tous ces arbres associés à un polynôme générique de degré fixé, et ensuite pour un arbre de lemniscate résultant d’un polynôme aléatoire avec des zéros indépendants et identiquement distribués. D’un point de vue plus général, ces résultats constituent un premier pas vers un traitement probabiliste (dans un cadre spécialisé) du programme d’Arnold consistant à énumérer les fonctions algébriques de Morse.

## 1. Introduction

Hilbert's sixteenth problem asks for an investigation of the topology of real algebraic curves and hypersurfaces. An extension of this program, promoted by V.I. Arnold [1], is to study the possible equivalence classes of graphs of generic polynomials up to diffeomorphism of the domain and range. Thus, rather than considering a single level set of a polynomial, Arnold's problem is concerned with the whole landscape given by its graph.

The special case of classifying graphs arising from the modulus  $|p(z)|$  of a generic complex polynomial  $p$  was solved by F. Catanese and M. Paluszny in [9].<sup>(1)</sup> They enumerated all possible equivalence classes by establishing a one-to-one correspondence with the combinatorial class of labeled, increasing, nonplane, binary trees.

Motivated by recent studies on the topology of *random* real algebraic varieties [14, 15, 17, 16, 21, 22, 24, 25, 26, 29], we investigate a probabilistic version of Arnold's problem by studying the topological properties of the landscape generated by the modulus of a random polynomial. We focus specifically on the statistical properties of a certain binary tree, the lemniscate tree defined below, of a random polynomial with independent identically distributed (i.i.d.) zeros drawn from a fixed probability measure on the Riemann sphere.

In this setting, the typical binary trees do not resemble the "combinatorial baseline" provided by sampling uniformly from the combinatorial class of all such trees (see Theorem 1.2 and compare with Theorem 1.1). Namely, the random tree associated to a random polynomial with i.i.d. zeros typically has very little branching (with probability converging to one in the limit of large degree, a shrinking portion of the nodes have two children).

### 1.1. Lemniscate trees

As in [9], we will call a polynomial  $p \in \mathbb{C}[z]$  of degree  $n + 1$  *lemniscate generic* (or simply *generic*) if  $p'$  has  $n$  distinct zeros  $w_1, \dots, w_n$  such that for each  $1 \leq i \leq n$ ,  $p(w_i) \neq 0$  and such that  $|p(w_i)| = |p(w_j)|$  if and only if  $i = j$ .<sup>(2)</sup> To such a polynomial one can associate a rooted, nonplane, binary

<sup>(1)</sup> This problem fits into Arnold's setting of real polynomials if we equivalently consider the square of the modulus and notice that  $p(z)\overline{p(z)}$  has real coefficients as a polynomial in  $x$  and  $y$ .

<sup>(2)</sup> The complement of the set of lemniscate generic polynomials forms a set of codimension one in the parameter space; as a result, in many models of random polynomials (including the ones studied in Section 3 of this paper) the condition of being lemniscate generic holds with probability one.

tree  $LT(p)$  with  $n$  vertices, whose vertices are bijectively labeled with the integers from 1 to  $n$  such that the labels increase along any path oriented away from the root. We call  $LT(p)$  the *lemniscate tree* associated to  $p$ .  $LT(p)$  encodes the topology of the graph of  $|p(z)|$  (or equivalently  $|p(z)|^2$  which can be viewed as a Morse function). Its vertices correspond to the  $n$  generically distinct critical points of  $p$ . To construct its edges consider for each critical point  $w$  of  $p$  the connected component

$$\Gamma_w \subseteq \{z \in \mathbb{C} \mid |p(z)| = |p(w)|\}$$

of the level set of  $|p|$  that contains  $w$ . The curve  $\Gamma_w$ , referred to as a *small lemniscate* in [9], is generically a bouquet of two circles with the self-crossing occurring precisely at  $w$ . The root vertex of  $LT(p)$  corresponds to the critical point with the largest value of  $|p|$ , which is generically unique. The descendants of the root are defined inductively as follows. Fix a vertex in  $LT(p)$  corresponding to a critical point  $w$  of  $p$ . Its children are the vertices in  $LT(p)$  corresponding critical points  $w'$  of  $p$  for which  $|p(w')|$  is the largest among all critical points whose singular lemniscates  $\Gamma_{w'}$  are surrounded by the same petal of  $\Gamma_w$ . Generically,  $\Gamma_w$  has two petals and each one gives rise to at most one descendent since the values of  $|p|$  at distinct critical points are generically distinct. Each vertex in  $LT(p)$  therefore generically has zero, one, or two children. We refer the reader to [9] for more details.

A simple example illustrating a polynomial and its corresponding lemniscate tree is provided in Figure 1.1. The left panel in Figure 1.1 displays all the singular level sets (each of which may include smooth components in addition to the singular component) for the modulus of a degree five polynomial, and the right panel shows the its lemniscate tree. To get a sense of what high-degree lemniscates can look like, consider Figure 1.3, where polynomials are generated by sampling zeros i.i.d. uniformly from the unit disk. A highly non-generic lemniscate, the so-called *Erdős lemniscate*  $\{z \in \mathbb{C} : |z^N - 1| = 1\}$  with  $N = 8$ , is shown in Figure 1.2. Figure 1.4 shows the singular lemniscates of a random polynomial generated by a linear combination of Chebyshev polynomials with Gaussian coefficients (see Section 4 for further discussion of this model).

While the lemniscate trees defined above are undirected, it will be convenient for us to adopt a term associated with directed graphs. We may impose an implicit direction on the edges of the trees so that each edge is oriented away from the root (such an oriented tree is properly called an *out-arborescence* but there will be no confusion here). In this context the number of children a vertex has is its *outdegree*, defined to be the number of directed edges emanating from it.

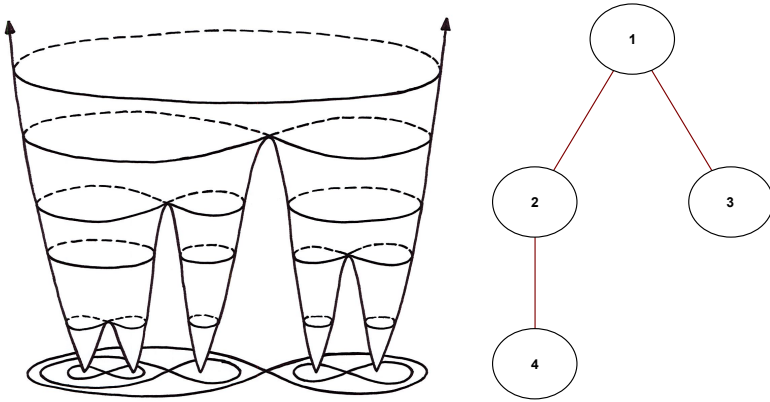


Figure 1.1. Left: an example of a landscape (artist's rendition) generated by a polynomial with five zeros, along with the projection of the singular component of each critical level. Right: the associated lemniscate tree (each node corresponds to a singular component). The tree can be constructed using the nesting structure of the singular components along with the ordering of heights of critical values.

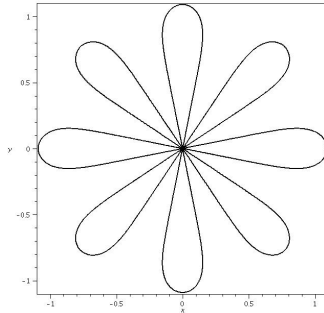


Figure 1.2. The non-generic lemniscate  $\{z \in \mathbb{C} : |z^8 - 1| = 1\}$ .

### 1.2. Random lemniscate trees

In this section we state our main results on the branching in random lemniscate trees. For each  $n \geq 1$  we define  $LT_n$  to be the set of all lemniscate trees on  $n$  vertices. That is,  $LT_n$  is the set of all rooted, nonplane, binary trees, with vertices bijectively labeled with the integers from 1 to  $n$  such that the labels increase along every path oriented away from the root. As mentioned above,  $LT_n$  was shown in [9] to be the space of possible

lemniscate trees for generic polynomials of degree  $n + 1$  in one complex variable. For every  $n \geq 1$ , the space  $LT_n$  is finite, and our first result concerns the branching structure of a tree sampled uniformly at random from  $LT_n$ .

**THEOREM 1.1.** — *Let  $T_n \in LT_n$  be a lemniscate tree of size  $n$  sampled uniformly at random, and let  $X_n$  denote the number of vertices of outdegree two in  $T_n$ . Write  $\mu_n$  for its mean and  $\sigma_n$  for its standard deviation. Then as  $n \rightarrow \infty$*

$$\mu_n = \left(1 - \frac{2}{\pi}\right)n + O(1),$$

and

$$\sigma_n^2 = \left(\frac{4}{\pi^2} + \frac{2}{\pi} - 1\right)n + O(1).$$

Moreover, the rescaled random variable  $\sigma_n^{-1}(X_n - \mu_n)$  converges in distribution to a standard Gaussian random variable as  $n \rightarrow \infty$ .

We note that the asymptotic for the mean follows from the asymptotic for the mean number of leaves which was computed recently in [5] (cf. [4, 7, 6]), where the same class of trees was referred to as 1-2 trees. By a standard application of Chebyshev's inequality, we see that Theorem 1.1 implies that the number of nodes of outdegree two is concentrated about its mean. Indeed, choosing  $0 < \alpha < 1/2$  we have

$$P(|X_n - \mu_n| > n^{\alpha+1/2}) \leq \frac{\sigma_n^2}{n^{\alpha+1/2}} = O(n^{\alpha-1/2}) = o(1), \quad \text{as } n \rightarrow \infty.$$

Therefore, for a uniformly randomly sampled  $T_n \in LT_n$ , one expects a constant proportion of its vertices to have two children. Our next result concerns the number of outdegree 2 nodes in the lemniscate tree of a random polynomial. Formally, we equip the space of polynomials of degree  $N$  with a measure under which zeros are chosen i.i.d. on the Riemann sphere, and push forward this measure to  $LT_{N-1}$  under the map that associates to a generic polynomial its lemniscate tree.

**THEOREM 1.2.** — *Let  $p_N$  be a random polynomial of degree  $N$  whose zeros are drawn i.i.d. from a fixed probability measure  $\mu$  on  $S^2$  that has a bounded density with respect to the uniform (Haar) measure. Then for every  $\epsilon > 0$  there exists  $C_\epsilon$  so that the number  $Y_N$  of nodes of outdegree two in the lemniscate tree associated to  $p_N$  satisfies*

$$\mathbb{E}Y_N \leq C_\epsilon N^{\frac{1}{2}+\epsilon}.$$

Although Theorem 1.2 does not give variance estimates and asymptotic normality as in Theorem 1.1, it does show that in contrast to sampling uniformly from  $LT_N$  the lemniscate tree of a random polynomial has almost all nodes with outdegree at most one. It also provides a weak concentration inequality for the random variables  $Y_N$ . Namely, for any  $\varepsilon_0 > 0$

$$P\left(Y_N > N^{\frac{1}{2}+\varepsilon_0}\right) \leq \frac{\mathbb{E}Y_N}{N^{\frac{1}{2}+\varepsilon_0}} = o(1), \quad \text{as } N \rightarrow \infty.$$

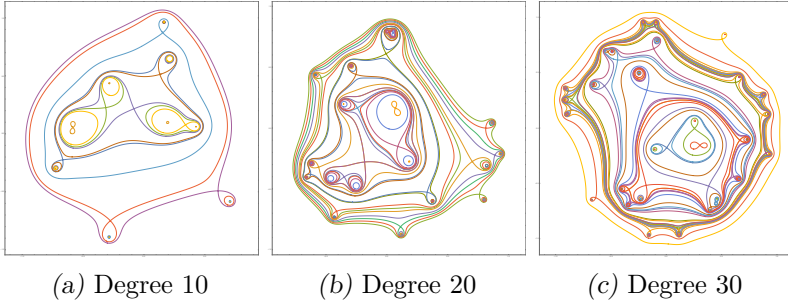


Figure 1.3. Lemniscates associated to random polynomials generated by sampling i.i.d. zeros distributed uniformly on the unit disk. For each of the three polynomials sampled, we have plotted (using Mathematica) each of the lemniscates that passes through a critical point. One observes a trend: most of the singular components have one large petal (surrounding additional singular components) and one small petal that does not surround any singular components. Note that only one of the connected components in each singular level set is singular (the rest of the components at that same level are smooth ovals).

This sparse branching for the lemniscate tree of a random polynomial is closely related to the pairing of zeros and critical points for random polynomials studied by the second author [18, 19, 20] and taken up in [28] as well. These articles roughly show that for the random polynomials we consider, each zero of  $p$  has, with high probability, a paired critical point in its  $1/N$  neighborhood. As we show in the proof of Theorem 1.2, when such a pairing occurs, the singular lemniscate  $\Gamma_w$  passing through the critical point  $w$  of  $p$  that is paired to a zero  $z$  is likely to have a small petal of diameter on the order of  $1/N$  surrounding  $z$  and no other zeros (and hence no other singular lemniscates), causing the corresponding vertex to have outdegree at most 1.

A useful heuristic for understanding the critical point pairing (which we combined with a topological argument in order to prove Theorem 1.2) is

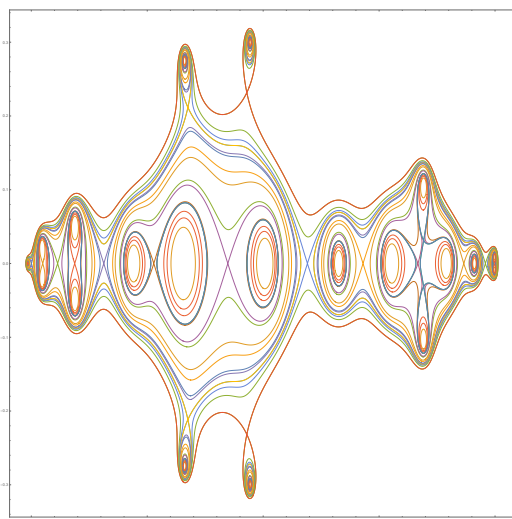


Figure 1.4. Lemniscates associated to a random linear combination of Chebyshev polynomials with Gaussian coefficients. Degree  $N = 20$ . This example is not lemniscate generic (since we see multiple critical points on a single level set). However, this model has the interesting feature that it seems to generate trees typically having many branches. See Section 4.

in terms of electrostatics on  $S^2$ , where critical points of  $p$  are viewed as equilibria of the field generated by a logarithmic potential with positively charged point particles at the zeros and negatively charged particles at the poles, counted with multiplicity.

The contribution to the electric field from the high-order pole at infinity (which is best understood after changing coordinates by  $z = 1/w$ ) is balanced by the electric field from an individual zero in a neighborhood with radius of order  $1/N$ . In this neighborhood, the additional influence of other zeros of  $p$  is typically of lower order, causing an almost deterministic pairing of zeros and critical points. We refer the reader to Section 3 below and to [20, §1] for more details.

*Remark 1.3.* — There is a fair amount of “universality” expressed in Theorem 1.2 in that the distribution  $\mu$  is rather arbitrary. What if the polynomial is instead sampled using random coefficients in front of some choice of basis? Based on simulations, the lemniscate trees again seem to have a shrinking portion of nodes with two children in a wide variety of such models including most of the well-studied Gaussian models (the Kostlan



model, the Weyl model, the Kac model). In fact, the only exception we observed was a model based on Chebyshev polynomials (see the empirical evidence presented below in the last section). In another direction, one might consider randomizing the construction of polynomial “fireworks” described in [13, §4] in order to produce polynomials whose trees have many branches.

*Remark 1.4.* — As a future direction of study it seems natural to investigate random *rational* functions on the Riemann sphere. A combinatorial scheme for classifying associated topological types was developed in [2]. What positive statements can one make on the typical topological type? The results in [22], investigating a fixed level set of a random rational function (defined as the ratio of two random polynomials from the Kostlan ensemble), may lead to some insight in this direction. However, we generally anticipate the case of rational functions to have a much different flavor than the case of polynomials; not only is the underlying combinatorial class more complicated, but there is no longer a “polarization” caused by having a high-order pole at infinity.

*Remark 1.5.* — Another natural direction of study, returning to Arnold’s problem mentioned at the beginning of the introduction, would be to investigate the topological type of a random homogeneous polynomial in projective space. The underlying classification problem in this case is still unsolved; L. Nicolaescu classified generic Morse functions on the 2-sphere [27] and enumerated them in terms of their number of critical points, but it is not known which types can be realized within each space of polynomials of given degree (and even less is known in more than two variables) [1]. At this stage, we suggest investigating a coarser structure, such as the so-called “merge tree” [11, §VII.1], associated to the graph of a random real homogeneous polynomial of degree  $d$  in  $n + 1$  variables (while pursuing asymptotic estimates as  $d \rightarrow \infty$  for statistics defined on the merge tree).

### 1.3. Outline of the paper

The Gaussian limit law stated in Theorem 1.1 will be established using *perturbed singularity analysis*, a method from analytic combinatorics. Specifically, in Section 2, we will apply a result from [12] to a bivariate generating function that was derived in [8]. We prove Theorem 1.2 in Section 3 by establishing a prevalence of small lemniscate petals adapting the method from [20] for studying pairing between zeros and critical points of

random polynomials. In Section 4, we present some empirical results concerning a certain model of random polynomials for which the lemniscate trees appear to have on average asymptotically one third of their nodes being of outdegree two.

### Acknowledgements

The authors would like to thank Alexandra Milbrand for creating the picture of the landscape of level sets in figure 1.1. We would like to thank Guilherme Silva for pointing out the reference [10] to a relevant ensemble of random matrices. We also thank the anonymous referee for a careful reading and helpful comments.

## 2. Sampling uniformly from the combinatorial class: proof of Theorem 1.1

Let  $a_{n,k}$  denote the number of lemniscate trees of size  $n$  with  $k$  nodes of outdegree two, and consider the bivariate generating function

$$F(z, u) = \sum_{n,k \geq 0} \frac{a_{n,k}}{n!} u^k z^n.$$

In [8], an explicit formula for the function  $F(z, u)$  is derived by showing that  $F$  satisfies a first-order PDE that can be solved explicitly using the method of characteristics. This results in the following analytic description in terms of elementary functions

$$(2.1) \quad F(z, u) = \left[ \cosh\left(\frac{z}{2}\sqrt{1-2u}\right) - \frac{\sinh\left(\frac{z}{2}\sqrt{1-2u}\right)}{\sqrt{1-2u}} \right]^{-2}.$$

There is a well-established theory for deriving probabilistic results from bivariate generating functions such as  $F(z, u)$ . For a detailed overview, see the authoritative text [12, Ch. IX] by Ph. Flajolet and R. Sedgewick; here we briefly review the connection in the current context. The basic link is that we arrive at the so-called *probability generating function* by considering a normalized coefficient extraction involving  $F(z, u)$ . Namely, using  $[z^n]$  to denote the operation of extracting the  $z^n$ -coefficient, the univariate polynomial in  $u$ , given by

$$p_n(u) = \frac{[z^n]F(z, u)}{[z^n]F(z, 1)},$$

is the probability generating function for the random variable  $X_n$  defined (as in the statement of Theorem 1.1) as the number of nodes of outdegree two in a random lemniscate tree of size  $n$ . That is, if a lemniscate tree of size  $n$  is sampled uniformly at random, the probability that it has  $k$  nodes of outdegree two is given by the coefficient of  $u^k$  in  $p_n(u)$ . From this, one can easily compute the mean and variance using simple operations. Furthermore, a more detailed complex analysis of the singularity structure of bivariate generating functions such as  $F(z, u)$  can be used to establish probabilistic limit laws.

Concerning the case at hand, viewing  $u$  as a complex parameter, the function  $F(z, u)$  is amenable to perturbed singularity analysis and falls under the “movable singularities schema” described in [12]; as  $u$  varies in a neighborhood of  $u = 1$ , the location of the (nearest to the origin) singularity of  $F(z, u)$  moves while the nature of this singularity is preserved. This allows us to establish a Gaussian limit law by apply the following result restated from [12, Thm. IX.12].

**THEOREM 2.1.** — *Let  $F(z, u)$  be a function that is bivariate analytic at  $(z, u) = (0, 1)$  and has non-negative coefficients. Assume the following conditions hold:*

- (i) *Analytic perturbation: there exist three functions  $A, B, C$ , analytic in a domain  $\mathcal{D} = \{|z| \leq r\} \times \{|u - 1| < \varepsilon\}$ , such that the following representation holds in some neighborhood of  $(0, 1)$ , with  $\alpha \notin \mathbb{Z}_{\leq 0}$ ,*

$$F(z, u) = A(z, u) + B(z, u)C(z, u)^{-\alpha}.$$

*Furthermore, in  $|z| \leq r$ , there exists a unique root  $\rho_1$  of the equation  $C(z, 1) = 0$ , this root is simple, and  $B(\rho_1, 1) \neq 0$ .*

- (ii) *Non-degeneracy: one has  $\partial_z C(\rho_1, 1) \cdot \partial_u C(\rho_1, 1) \neq 0$ , ensuring the existence of a non-constant analytic function  $\rho(u)$  near  $u = 1$ , such that  $C(\rho(u), u) = 0$  and  $\rho(1) = \rho_1$ .*
- (iii) *Variability: one has*

$$v(\beta) := \frac{\beta''(1)}{\beta(1)} + \frac{\beta'(1)}{\beta(1)} - \left( \frac{\beta'(1)}{\beta(1)} \right)^2 \neq 0,$$

*where  $\beta(u) = \rho(1)\rho(u)^{-1}$ .*

*Then, the random variable with probability generating function*

$$p_n(u) = \frac{[z^n]F(z, u)}{[z^n]F(z, 1)}$$

*converges in distribution (after standardization) to a Gaussian random variable with a speed of convergence  $O(n^{-1/2})$ .*

*Verifying condition (i).* — Let  $G(z, u) = \cosh\left(\frac{z}{2}\sqrt{1-2u}\right) - \frac{\sinh\left(\frac{z}{2}\sqrt{1-2u}\right)}{\sqrt{1-2u}}$  so that we have  $F(z, u) = G(z, u)^{-2}$ . Note that  $G(z, u)$  is an entire function of  $z$  for each fixed  $u \in \mathbb{C}$ , and is non-constant for  $u \neq \frac{1}{2}$ . Thus, for  $u \neq \frac{1}{2}$ ,  $F(z, u)$  is meromorphic with poles at the zeros of  $G(z, u)$  and no other singularities. First setting  $u = 1$ , we find that the zeros of  $G(z, 1)$  are at  $z = \frac{\pi}{2} + 2\pi i k, k \in \mathbb{Z}$ . Among these,  $\rho_1 = \frac{\pi}{2}$  is nearest to the origin. We compute  $\partial_z G(\rho_1, 1) = \frac{-\sqrt{2}}{2} \neq 0$ , which shows that  $\rho_1$  is a simple root. This completes the verification of condition (i) in Theorem 2.1, where  $A = 1, B = 1$  are taken to be constant,  $\alpha = 2$ , and  $C(z, u) = G(z, u)$ .

*Verifying condition (ii).* — Having shown above that  $\partial_z G(\rho_1, 1) \neq 0$ , we only need to check that  $\partial_u G(\rho_1, 1) \neq 0$ . We find  $\partial_u G(\rho_1, 1) = \frac{\sqrt{2}}{4}(2 - \pi)$ . This verifies condition (ii), where the desired function  $\rho(u)$  is guaranteed to exist by the implicit function theorem. Furthermore, we can describe  $\rho(u)$  explicitly by solving for  $z$  in  $G(z, u) = 0$ :

$$\rho(u) = \frac{2}{\sqrt{1-2u}} \tanh^{-1}(\sqrt{1-2u}) = \frac{1}{\sqrt{1-2u}} \log\left(\frac{1 + \sqrt{1-2u}}{1 - \sqrt{1-2u}}\right),$$

where we choose the principal branch for the logarithm so that  $\rho(1) = \rho_1 = \pi/2$ . For  $|u - 1|$  sufficiently small,  $\sqrt{1-2u}$  is near  $i$ , ensuring analyticity of the function  $\rho(u)$ .

*Verifying condition (iii).* — Let  $\beta(u) = \rho(1)/\rho(u)$ . Then

$$\beta'(u) = -\frac{\pi}{2}\rho(u)^{-2}\rho'(u),$$

and

$$\beta''(u) = -\frac{\pi}{2}\left(-2\rho(u)^{-3}(\rho'(u))^2 + \rho(u)^{-2}\rho''(u)\right).$$

Thus,

$$\begin{aligned} v(\beta) &= \frac{\beta''(1)}{\beta(1)} + \frac{\beta'(1)}{\beta(1)} - \left(\frac{\beta'(1)}{\beta(1)}\right)^2 \\ &= \left(\frac{8}{\pi^2} - 1\right) + \left(1 - \frac{2}{\pi}\right) - \left(1 - \frac{2}{\pi}\right)^2 \\ &= \frac{4}{\pi^2} + \frac{2}{\pi} - 1 \neq 0. \end{aligned}$$

We conclude that Theorem 2.1 applies, and the random variable  $N_2(T_n)$  after rescaling converges in distribution to a Gaussian variable with a speed of convergence  $O(n^{-1/2})$ .

*Mean and variance.* — As pointed out in [12] in the remarks after the proof of Theorem IX.12, the mean  $\mu_n$  and variance  $\sigma_n^2$  are given by

$$\mu_n = \frac{\beta'(1)}{\beta(1)}n + O(1) = \left(1 - \frac{2}{\pi}\right)n + O(1),$$

and

$$\sigma_n^2 = v(\beta) \cdot n + O(1) = \left(\frac{4}{\pi^2} + \frac{2}{\pi} - 1\right)n + O(1).$$

### 3. Proof of Theorem 1.2

For a polynomial  $p$  of one variable we define

$$\text{Zeros}(p) := \{\xi \in S^2 \text{ s.t. } p(\xi) = 0\}, \quad \text{Crits}(p) := \{w \in S^2 \text{ s.t. } dp(w) = 0\}.$$

Instead of working the “usual” holomorphic coordinate  $S^2 \setminus \{\infty\} \rightarrow \mathbb{C}$ , it will be more convenient to perform our computations in the coordinate  $S^2 \setminus \{0\} \rightarrow \mathbb{C}$  centered at the point at infinity (cf. [20, §1]). That is, we first choose  $\xi_1, \dots, \xi_N \in S^2$  i.i.d. from the probability measure  $\mu$ . Up to multiplication by a non-zero complex number, this determines a unique meromorphic map  $p_N$  from  $S^2$  to  $\mathbb{C} \cup \{\infty\}$  with a pole of order  $N$  at infinity and zeros at the  $\xi_j$ 's. Since multiplication by a non-zero constant does not affect zeros, critical points, and relative sizes  $|p_N(z_1)|/|p_N(z_2)|$  of the modulus of  $p$  for any  $z_1, z_2 \in S^2$ , the zeros  $\xi_1, \dots, \xi_N \in S^2$  uniquely determine the lemniscate tree of  $p_N$ . We investigate the structure of this lemniscate tree using the coordinate  $w : S^2 \setminus \{0\} \rightarrow \mathbb{C}$  centered at infinity, in which  $p_N$  takes the form

$$p_N(w) := \frac{1}{w^N} \prod_{j=1}^N (w - \xi_j),$$

where the i.i.d. zeros  $\xi_j$  are now written in the coordinate  $w$  (we abuse notation slightly and continue to denote them by  $\xi_j$  in this coordinate). Let us emphasize that whenever a condition like  $|\xi| \leq N^\Delta$  appears below for some  $\xi \in S^2$ , the quantity  $|\xi|$  is computed in this system of coordinates. In particular, denoting by  $\zeta = 1/\xi$  the image of  $\xi$  in the usual coordinates centered at 0, our condition  $|\xi| \leq N^\Delta$  is the same as  $|\zeta| \geq N^{-\Delta}$ . Associated to each  $w \in \text{Crits}(p_N)$  is the singular component  $\Gamma_w$  of the lemniscate

$$\Lambda_w := \{z \in S^2 \text{ s.t. } |p_N(z)| = |p_N(w)|\},$$

that passes through  $w$ . That is, among the connected components of  $\Lambda_w$ , we define  $\Gamma_w$  to be the one that contains  $w$ . For a *generic* polynomial (a

condition that holds with probability one in our model), there are  $N - 1$  distinct singular lemniscates (one passing through each critical point), each having a unique singular component that is topologically a bouquet of two circles. We call these two circles *the petals* of  $\Gamma_w$ . For the arguments below we fix an auxiliary parameter  $r \gg 1$  such that

$$(3.1) \quad \arg \left( 1 + \frac{1}{r} e^{i\theta} \right) \in \left( -\frac{1}{10}, \frac{1}{10} \right), \quad \forall \theta \in [0, 2\pi].$$

We study the behavior of the lemniscate tree of  $p_N$  by considering for each  $\xi \in \text{Zeros}(p_N)$  the event

$$S_{\xi,N} := \left\{ \exists! w \in \text{Crits}(p_N) \left| \begin{array}{l} |\xi - w| < \frac{r|\xi|}{N} \text{ and at least one petal of } \Gamma_w \\ \text{is contained in the disk of radius } \frac{4r|\xi|}{N} \\ \text{centered at } \xi \end{array} \right. \right\}.$$

When the event  $S_{\xi,N}$  occurs, we will say that  $\Gamma_w$  has a *small petal surrounding*  $\xi$ , and we refer to  $w$  as the *paired critical point* of  $\xi$ . We also consider the events

$$B_{\xi,N} := \left\{ |\xi - \xi'| > \frac{4r|\xi|}{N}, \quad \forall \xi' \in \text{Zeros}(p_N) \setminus \{\xi\} \right\}.$$

To prove Theorem 1.2, we begin with the following observation.

LEMMA 3.1. — *We have*

$$(3.2) \quad \#\{\text{vertices in } LT(p_N) \text{ with at most one child}\} \geq \sum_{\xi \in \text{Zeros}(p_N)} \mathbb{1}_{S_{\xi,N} \cap B_{\xi,N}},$$

where  $\mathbb{1}_S$  denotes the indicator function of the event  $S$ .

*Proof.* — Observe that if the events  $S_{N,\xi} \cap B_{N,\xi}$  and  $S_{N,\xi'} \cap B_{N,\xi'}$  occur for some zeros  $\xi \neq \xi'$ , then the corresponding paired critical points are also distinct since the spacing of zeros ensured by  $B_{\xi,N}$  and  $B_{\xi',N}$  is larger than the sum of the distances between the zeros to their paired critical points given by  $S_{\xi,N}$  and  $S_{\xi',N}$ . Moreover, when the event  $S_{\xi,N} \cap B_{\xi,N}$  occurs, the vertex in the lemniscate tree of  $p_N$  that corresponds to the critical point  $w$  paired to  $\xi$  has outdegree at most 1. Indeed, one of its petals surrounds only one zero, namely  $\xi$ . To complete the proof, we recall a theorem of Titchmarsh ([30, §3.55]), which says that inside a connected component of a lemniscate for an analytic function the number of critical points is one less than the number of zeros. The petal surrounding  $\xi$  has inside one zero and therefore does not surround any critical points. Hence, the vertex in the lemniscate tree of  $p_N$  that corresponds to  $w$  has at most one child.  $\square$

The preceding Lemma shows that

$$(3.3) \quad \mathbb{E}\#\{\text{vertices in } LT(p_N) \text{ with at most one child}\} \geq N \cdot \mathbb{P}(S_{\xi,N} \cap B_{\xi,N}).$$

To obtain a lower bound for the probability of  $S_{\xi,N} \cap B_{\xi,N}$ , note that for any  $\Delta \in [0, 1/4)$  and any  $\xi \in \text{Zeros}(p_N)$  we have

$$\mathbb{P}(|\xi| \leq N^\Delta) = 1 + O(N^{-2\Delta})$$

since the measure  $\mu$  assigns to a ball of radius  $N^{-\Delta}$  centered at any point on  $S^2$  (in particular at 0) a mass on the order of its volume. Using that  $\mu$  has a bounded density with respect to the uniform measure on  $S^2$ , we have

$$\begin{aligned} \mathbb{P}(B_{\xi,N}) &= \mathbb{P}(B_{\xi,N} \mid |\xi| \leq N^\Delta) (1 + O(N^{-2\Delta})) \\ &= \left[ 1 - (N - 1) \int_{|\xi| \leq N^\Delta} \mu \left( \left\{ \zeta \mid |\xi - \zeta| < \frac{4r|\xi|}{N} \right\} \right) d\mu(\xi) \right] (1 + O(N^{-2\Delta})) \\ &= [1 - O(N^{-1+2\Delta})] (1 + O(N^{-2\Delta})) \\ &= 1 + O(N^{-2\Delta}) + O(N^{-1+2\Delta}). \end{aligned}$$

Therefore, since for  $\Delta \in [0, 1/4)$ , we have  $-2\Delta > -1 + 2\Delta$ , we find that

$$(3.4) \quad \begin{aligned} \mathbb{P}(S_{\xi,N} \cap B_{\xi,N}) &= \mathbb{P}(S_{\xi,N}) + O(N^{-2\Delta}) \\ &\geq \int_{|\xi| \leq N^\Delta} \mathbb{P}(S_{\xi,N} \mid \xi) d\mu(\xi) + O(N^{-2\Delta}), \end{aligned}$$

where the notation in the last integral is that we've conditioned on the position of  $\xi$ . We now fix  $\Delta \in [0, 1/4)$ , a deterministic sequence  $\xi = \xi_N$  with  $|\xi| \leq N^\Delta$ , and consider the random polynomials

$$p_{\xi,N}(w) := \frac{1}{w^N} (w - \xi) \prod_{j=1}^{N-1} (w - \xi_j),$$

conditioned to have a zero at  $\xi$  and with  $\xi_j$  drawn i.i.d. from  $\mu$  for  $j = 1, \dots, N - 1$ . We slightly abuse notation and continue to write  $S_{\xi,N}$  for the event that (the fixed zero)  $\xi$  has a paired critical point  $w_{\xi,N}$  with a small petal surrounding  $\xi$ , so the conditional probability appearing in the integrand in (3.4) is henceforth simply denoted as  $\mathbb{P}(S_{\xi,N})$ . Theorem 1.2 follows from (3.4) once we show that there exists  $N_0 \geq 1$  and  $C_\Delta > 0$  so

that for all  $N \geq N_0$

$$(3.5) \quad \inf_{|\xi| \leq N^\Delta} \mathbb{P}(S_{\xi,N}) \geq 1 - C_\Delta N^{-2\Delta}.$$

To show (3.5), we revisit the proof of the main theorem in [20]. To state the precise estimate we will use, we set some notation. Critical points of  $p_{N,\xi}$  are solutions to  $E_N(w) = 0$ , where

$$(3.6) \quad E_N(w) = d \log p_{N,\xi}(w) = -\frac{N}{w} + \frac{1}{w - \xi} + \sum_{j=1}^{N-1} \frac{1}{w - \xi_j}.$$

As in [20, §4], observe that

$$(3.7) \quad \mathbb{E}E_N(w) = -\frac{N}{w} + \frac{1}{w - \xi} + (N - 1) \int_{\mathbb{C}} \frac{d\mu(z)}{w - z},$$

and

$$(3.8) \quad \tilde{E}_N(w) := E_N(w) - \mathbb{E}E_N(w) = \sum_{j=1}^{N-1} \frac{1}{w - \xi_j}.$$

In the computations below, the Cauchy–Stieltjes transform  $\int_{\mathbb{C}} \frac{d\mu(z)}{w - z}$  appearing in (3.7) plays no significant role (it only shifts the locations of critical points in a deterministic way so that (3.9) below has an additional deterministic  $1/N$  correction). Hence, we will assume that it is identically 0 (i.e. we reduce to the case when  $\mu$  is the uniform measure on  $S^2$ ). For each  $\Delta \in [0, 1/4)$  and all  $\xi$  with  $|\xi| \leq N^\Delta$ , the average critical point equation  $\mathbb{E}E_N(w) = 0$  has a unique solution

$$(3.9) \quad w_{\xi,N} := \xi \left(1 - \frac{1}{N}\right)^{-1}$$

near  $\xi$ . Note that

$$|w_{\xi,N} - \xi| = \frac{|\xi|}{N - 1}.$$

We will argue in Section 3.1 below that the technique in [20] gives the following proposition.

PROPOSITION 3.2. — Fix  $\Delta \in (0, 1/4)$ . For each  $\xi$  let  $D_{N,\xi}$  denote the disk of radius  $4r|\xi|/N$  centered at  $\xi$ . There exists  $\gamma > 0$  and a constant  $C_\Delta$  so that the event

$$X_{N,\xi,\Delta} = \left\{ \sup_{w \in D_{N,\xi}} \left| \tilde{E}_N(w) \right| \leq N^{1-\gamma} / |\xi| \right\}$$



occurs with high probability:

$$(3.10) \quad \inf_{|\xi| \leq N^\Delta} \mathbb{P}(X_{N,\xi,\Delta}) \geq 1 - C_\Delta \cdot N^{-2\Delta}.$$

Assuming Proposition 3.2 for the moment, we complete the proof of (3.5) and hence of Theorem 1.2 by showing that the event  $X_{N,\xi,\Delta}$  (or more precisely  $X_{N,\xi,\Delta} \cap B_{N,\xi}$ ), whose probability is estimated in (3.10), is contained in the event  $S_{\xi,N}$ . Suppose that  $X_{N,\xi,\Delta}$  occurs. Then, as in [20], by Rouché’s Theorem applied to  $E_N$ , there exists  $N_0$  so that for all  $N \geq N_0$  there is a unique  $w \in \text{Crit}(p_N)$  satisfying

$$|w - \xi| < r |\xi| / N$$

with probability at least  $1 - C_\Delta \cdot N^{-2\Delta}$ . Indeed, write  $\Gamma$  for the boundary of the disk of radius  $r |\xi| / N$  centered at  $\xi$ . Since  $r \gg 1$ , the curve  $\Gamma$  winds around  $w_{N,\xi}$  (defined in (3.9)) for all  $N$ . Moreover, by the triangle inequality,

$$\begin{aligned} \inf_{w \in \Gamma} |\mathbb{E}E_N(w)| &\geq \inf_{w \in \Gamma} \left| \frac{N}{|w|} - \frac{1}{|\xi - w|} \right| = \left| \frac{N}{|\xi| - \frac{|\xi|}{N-1}} - \frac{1}{r |\xi| / N} \right| \\ &= \frac{N}{|\xi|} \left( \frac{N-1}{N-2} - \frac{1}{r} \right). \end{aligned}$$

If  $X_{N,\xi,\Delta}$  happens, we also have

$$\sup_{w \in \Gamma} |\tilde{E}_N(w)| \leq \frac{N^{1-\gamma}}{|\xi|}.$$

Hence, on the event  $X_{N,\xi,\Delta} \cap B_{N,\xi}$  for which

$$\mathbb{P}(X_{N,\xi,\Delta} \cap B_{N,\xi}) = \mathbb{P}(X_{N,\xi,\Delta}) + O(N^{-1+2\Delta})$$

we find

$$\inf_{w \in \Gamma} |\mathbb{E}E_N(w)| > \sup_{w \in \Gamma} |\tilde{E}_N|.$$

We may therefore apply Rouché’s Theorem to conclude that  $E_N$  has exactly one zero (and hence  $p_{N,\xi}$  has exactly one critical point) in the interior of  $\Gamma$ . This is precisely the first condition in the definition of  $S_{N,\xi}$ .

To check that the small petal condition in the definition of  $S_{N,\xi}$  is also satisfied when  $X_{N,\xi,\Delta}$  occurs, let  $A_{N,\xi}$  denote the annulus centered at  $\xi$  with inner radius  $r |\xi| / N$  and outer radius  $4r |\xi| / N$  (recall that  $r$  was fixed by (3.1)). For simplicity, we will rotate our coordinates so that  $\xi$  lies on the positive real axis and consider the three regions in  $A_{N,\xi}$  (see Figure 3.1

below):

$$\begin{aligned} \Omega_1 &:= \{w \in A_{N,\xi} \mid \Re w < \xi - 3r\xi/N\}, \\ \Omega_2 &:= \{w \in A_{N,\xi} \mid \xi - 2r\xi/N < \Re w < \xi + 2r\xi/N\}, \\ \Omega_3 &:= \{w \in A_{N,\xi} \mid \Re w > \xi + 3r\xi/N\}. \end{aligned}$$

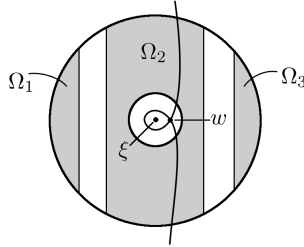


Figure 3.1. The lemniscate  $\{|p_N(z)| = |p_N(w)|\}$  passing through  $w$  has with high probability empty intersection with each of  $\Omega_1$  and  $\Omega_2$ . This forces one of the corresponding petals to be contained in  $D_{N,\xi}$  which implies that petal is empty (it contains no critical points in its interior).

We now argue that the event  $X_{N,\xi,\Delta}$  in (3.10) implies that the argument of  $E_N(w)$  is essentially deterministic (given by the argument of  $\mathbb{E}E_N(w)$  to leading order in  $N$ ) uniformly as  $w$  ranges over  $A_{N,\xi}$ . We parameterize points  $w \in A_{N,\xi}$  by writing  $w = \xi + \rho e^{i\theta}$  with  $r|\xi|/N < \rho < 4r|\xi|/N$ . Then (3.7) yields

$$(3.11) \quad \mathbb{E}E_N(w) = \frac{N}{|\xi|} \left( -\frac{1}{1 + \rho e^{i\theta}/|\xi|} + \frac{1}{N\rho e^{i\theta}/|\xi|} \right).$$

The first term inside the parentheses in (3.11) is  $-1 + O(1/N)$  since

$$\left| -\frac{1}{1 + \rho e^{i\theta}/|\xi|} + 1 \right| \leq \frac{4r}{N - 4r} = O\left(\frac{1}{N}\right),$$

while the second term satisfies

$$\frac{1}{4r} \leq \left| \frac{1}{N\rho e^{i\theta}/|\xi|} \right| \leq \frac{1}{r},$$

which we summarize by writing

$$\frac{1}{N\rho e^{i\theta}/|\xi|} = F(\rho) \frac{1}{r} e^{-i\theta},$$

and taking note that

$$\frac{1}{4} \leq F(\rho) \leq 1.$$

Therefore, Proposition 3.2 (along with the estimates above) shows that for each  $\Delta \in (0, 1/4)$  there exists  $\gamma > 0$  so that with probability at least  $1 - C_\Delta N^{-2\Delta}$ , we have

$$\begin{aligned} E_N(\xi + \rho e^{i\theta}) &= \mathbb{E}E_N(\xi + \rho e^{i\theta}) + \tilde{E}_N(\xi + \rho e^{i\theta}) \\ &= \frac{N}{|\xi|} \left( -1 + F(\rho) \frac{1}{r} e^{-i\theta} + O(N^{-\gamma}) \right), \end{aligned}$$

where the implied constant is independent of  $\rho, \theta, N, \xi$ . Thus, using the definition (3.1) of  $r$  we conclude when  $X_{N,\xi,\Delta}$  occurs, we also have

$$(3.12) \quad \arg(E_N(w)) \in \left( \pi - \frac{1}{5}, \pi + \frac{1}{5} \right), \quad \forall w \in A_{N,\xi}$$

for all  $N$  sufficiently large. Let us write

$$\begin{aligned} \partial_- \Omega_2 &= \{w \in A_{N,\xi} \mid \Re(w) = \xi - 2r\xi/N\}, \\ \partial_+ \Omega_2 &= \{w \in A_{N,\xi} \mid \Re(w) = \xi + 2r\xi/N\} \end{aligned}$$

for the left and right boundaries of  $\Omega_2$  and set

$$\zeta = \pi/2 - \arctan 8 > 0.$$

Note that the angle of any line segment joining a point on  $\partial_- \Omega_2$  to any point in  $\Omega_1$  lies in the interval  $(\pi/2 + \zeta, 3\pi/2 - \zeta)$ , so that it forms an acute angle with  $E_N(w)$  when  $X_{N,\xi,\Delta}$  happens by (3.12). Since  $E_N(w)$  has the same argument as the gradient of  $|p_{N,\xi}(w)|$ , this implies that the event  $X_{N,\xi,\Delta}$  entails that the directional derivative of  $|p_{N,\xi}(w)|$  along such a line segment is positive, and hence the value of  $|p_{N,\xi}(w)|$  in  $\Omega_1$  is strictly larger than its value on  $\partial_- \Omega_2$ . Similarly, the value of  $|p_{N,\xi}(w)|$  on the right boundary  $\partial_+ \Omega_2$  is strictly larger than its value throughout  $\Omega_3$ . This implies that a level curve of  $|p_{N,\xi}|$  that intersects  $\Omega_2$  cannot intersect  $\Omega_1$  or  $\Omega_3$  unless it leaves  $D_{N,\xi}$  through the set

$$S := \{|w - \xi| = 4r\xi/N\} \cap \Omega_2,$$

which consists of two circular arcs symmetric with respect to the real axis. As above, the event  $X_{N,\xi,\Delta}$  ensures that the argument of  $E_N$  is close to  $\pi$  and hence the restriction of  $|p_{N,\xi}|$  to each component of  $S$  is strictly monotone. Thus, any level curve of  $|p_{N,\xi}|$  can only cross each component of  $S$  once on the event  $X_{N,\xi,\Delta}$ . The singular component  $\Gamma_w$  (consisting of two petals joined at  $w$ ) of the lemniscate passing through the critical point  $w$  that is paired to  $\xi$  therefore crosses the boundary of  $D_{N,\xi}$  at most twice (one crossing for each component of  $S$ ). This implies whenever  $X_{N,\xi,\Delta}$  occurs, one of the petals is completely contained in  $D_{N,\xi}$ , and therefore it must be a small petal since  $D_{N,\xi}$  contains only one zero of  $p_N$ , see Figure 3.1.

This shows that  $X_{N,\xi,\Delta} \cap B_{N,\xi}$  implies  $S_{N,\xi}$  and yields (3.5), completing the proof of Theorem 1.2.  $\square$

### 3.1. Proof of Proposition 3.2

Fix  $\Delta \in (0, 1/4)$  and a sequence  $\xi = \xi(N)$  with  $|\xi| \leq N^\Delta$  (we remind the reader that  $|\xi|$  is measured in coordinates centered at  $\infty$ , and hence in terms of the original coordinates our assumption removes a disk of radius  $N^{-\Delta}$  centered at 0). In this section we explain how to modify the proof of Theorem 1 (specifically equation (4.2)) in [20] to prove Proposition 3.2. The argument from [20] was presented in several steps; below we explain the modifications needed at each step.

*Step 1.* — With  $w_\xi = w_{\xi,N}$  defined as in (3.9), we study  $\tilde{E}_N(w)$  by separately considering

$$\tilde{E}_N(w_\xi), \quad \text{and} \quad \tilde{E}_N(w) - \tilde{E}_N(w_\xi).$$

*Step 2.* — To understand

$$\tilde{E}_N(w) - \tilde{E}_N(w_\xi) = \sum_{j=1}^{N-1} \frac{w_\xi - w}{(w - \xi_j)(w_\xi - \xi_j)}$$

we first fix  $\delta \in (2\Delta, 1)$  and estimate the contribution from zeros far away from  $\xi$ :

$$(3.13) \quad \left| \sum_{|\xi - \xi_j| > N^{-1/2+\delta/2}} \frac{w_\xi - w}{(w - \xi_j)(w_\xi - \xi_j)} \right| \leq K_1 N^{1-\delta+\Delta},$$

for some  $K_1 > 0$ , where we've used that  $|w_\xi - w| \leq 4rN^{-1+\Delta}$  for  $w \in D_{N,\xi}$  and that  $|w - \xi_j|$  and  $|w_\xi - \xi_j|$  are each bounded from below by a positive constant times  $|\xi - \xi_j|$  (in fact  $|w - \xi_j|$ ,  $|w_\xi - \xi_j|$ , and  $|\xi - \xi_j|$  are asymptotically equivalent as  $N \rightarrow \infty$  as can be see from the triangle inequality). Hence, as long as  $\delta > 2\Delta$ , we find that the expression in (3.13) is (deterministically) bounded above by

$$N^{1-\epsilon-\Delta} \leq N^{1-\epsilon} / |\xi|$$

for some  $\epsilon > 0$ , as in the definition of the event  $X_{N,\xi}$  whose probability we seek to estimate.

*Step 3.* — Next, we control the contribution to  $\tilde{E}_N(w) - \tilde{E}_N(w_\xi)$  from zeros near  $\xi$  by successively subtracting and adding

$$\sum_{|\xi - \xi_j| \leq N^{-1/2 + \delta/2}} \frac{(w_\xi - w)^k}{(w_\xi - \xi_j)^{k+1}}$$

for  $k = 1, 2, \dots, L - 1$  and also recalling that  $|w_\xi - w| < 4rN^{-1 + \Delta}$ :

$$(3.14) \quad \left| \sum_{|\xi - \xi_j| \leq N^{-1/2 + \delta/2}} \frac{w_\xi - w}{(w_\xi - \xi_j)(w - \xi_j)} \right| \leq \left| \sum_{|\xi_j - \xi| \leq N^{-1/2 + \delta/2}} \frac{(w_\xi - w)^L}{(w_\xi - \xi_j)^L (w - \xi_j)} \right| + \sum_{k=1}^{L-1} c^k N^{(-1 + \Delta)k} \left| \sum_{|\xi - \xi_j| \leq N^{-1/2 + \delta/2}} \frac{1}{(w_\xi - \xi_j)^{k+1}} \right|,$$

where  $c$  is an absolute constant and  $L$  is any positive integer.

*Step 4.* — We control the two terms in (3.14) separately. To control the term containing the sum on  $k$ , we use [20, Lem. 2], which says that for every  $\eta \in (0, 1/2)$  there exists  $K = K(\eta) > 0$  so that

$$\mathbb{P} \left( \sum_{j=1}^{N-1} \frac{1}{|w_\xi - \xi_j|^2} > N^{2-2\eta} \right) \leq K \cdot N^{-1+2\eta} \log N.$$

Hence, taking  $\eta = \Delta + \epsilon/2$ , we find

$$\sum_{k=1}^{L-1} c^k N^{(-1 + \Delta)k} \left| \sum_{|\xi - \xi_j| \leq N^{-1/2 + \delta/2}} \frac{1}{(w_\xi - w)^{k+1}} \right| \leq \sum_{k=1}^{L-1} c^k N^{(-1 + \Delta)k} N^{(2-2\eta)\frac{k+1}{2}} \leq C_\epsilon N^{1 - \Delta - \epsilon} \leq C_\epsilon N^{1 - \epsilon} / |\xi|$$

with probability at least  $N^{-1 + 2\Delta + \epsilon}$ .

*Step 5.* — To bound the other term in (3.14), we use [20, Lem. 1], which says that for each  $\delta \in (0, 1/2)$  with probability at least  $1 - C_\delta N^{-\delta}$  there are no zeros with  $|\xi_j| > N^{1/2 + \delta/2}$  and at most  $N^{2\delta}$  zeros with  $|\xi_j| > N^{1/2 - \delta/2}$ . This allows us to write

$$\left| \sum_{|\xi_j| - \xi \leq N^{-1/2 + \delta/2}} \frac{(w_\xi - w)^L}{(w_\xi - \xi_j)^L (w - \xi_j)} \right| \leq N^{1 - (1 - 2\delta + L(1 - \Delta) - (L+1)(1/2 + \delta/2))}.$$

Hence, taking  $L$  sufficiently large, we find that if  $2\Delta < \frac{1}{2}$ , then the left hand side in the previous line can be bounded above by  $N^{-1+\epsilon}/|\xi|$  for all  $\epsilon > 0$  sufficiently small with probability at least  $1 - C_\Delta N^{-1+\Delta}$ . This completes the proof of Proposition 3.2.

#### 4. Random perturbation of a Chebyshev polynomial

In light of the results of the previous section, one may wonder whether there are any natural models of random polynomials that typically have some positive portion of the nodes in the corresponding tree having two children (thus resembling the previously established combinatorial baseline). As a possible candidate for such a model the authors considered random linear combinations of Chebyshev polynomials. More specifically, the class of polynomials considered were of the form  $p(z) = \sum_{k=0}^n a_k T_k(z)$ , where  $T_k$  is the Chebyshev polynomial (of the first kind) of degree  $k$ , and the coefficients are chosen independently with  $a_n \sim N(0, 1)$ . Linear combinations of orthogonal polynomials have been studied previously, including several varieties of Jacobi orthogonal polynomials [23]. The important property of Chebyshev polynomials (leading us to choose those as a basis) is that they each have critical values all with the same modulus. Figure 1.4 shows the family of singular level sets for such a polynomial. The lemniscates for this type of polynomial appear to exhibit a rich nesting structure. However, these polynomials are typically not lemniscate generic (due to complex conjugate pairs of critical points sharing the same critical value). Consequently, this model seems worthy of further investigation, but this will require first understanding an appropriate class of lemniscate trees.

We will investigate a modified (less organic, but more tractable) version of this model where the top degree Chebyshev polynomial gets most of the weight. Specifically, we consider randomly perturbed Chebyshev polynomials of the form  $T_n(z) + \frac{1}{n} \sum_{k=0}^{n-1} b_k T_k(z)$ , where the coefficients  $b_k$  are randomly and independently chosen to be 1 or  $-1$  with equal probability. These polynomials have all real roots and real critical points, which enables us to easily determine the corresponding lemniscate tree by the process described below.

Suppose that  $p$  is a lemniscate generic polynomial with real zeros and critical points. We construct a permutation as follows: we label the critical points with the integers 1 through  $\deg(p) - 1$ , starting with 1 for the critical point with largest critical value in magnitude, 2 for the critical point with second largest critical value in magnitude, and so on. Reading the labels

from left to right gives a permutation of the numbers 1 through  $\deg(p) - 1$ . Now, it is well known that the permutations on  $n$  letters are in one-to-one correspondence with the *increasing binary trees* of size  $n$  (see [12, p. 143] for example). These are plane, labeled, rooted trees in which every vertex has at most two children, where each child has a left or right orientation (even when it is the unique child of its parent), such that the labels along any path directed away from the root are increasing. We then construct the increasing binary tree corresponding to the permutation obtained from the polynomial. By “forgetting” its embedding in the plane we obtain the lemniscate tree associated to the singular level sets of the polynomial. One can even determine the number of nodes of outdegree 2 directly from the permutation by counting the number of descents which are immediately followed by ascents.

We apply this procedure to a number of polynomials in the following computer experiment: Table 4.1 gives the average value of  $N_2$  computed for a sample of 100 randomly perturbed Chebyshev polynomials of the same degree  $n$  for a number of different values of  $n$  ranging from 10 to 200. Linear regression yields a best fit line with equation  $N_2 = 0.3338n - 0.90803$  with  $R^2 = 0.9999$ , indicating that one should expect approximately a third of the vertices in the lemniscate tree for a perturbed Chebyshev polynomial to have outdegree two. This agrees with a heuristic of ignoring correlations in the randomly perturbed heights of critical values in the perturbed Chebyshev polynomial, which corresponds to the induced random permutation being sampled uniformly from the combinatorial class of permutations (it is known [3] that the average number of nodes of outdegree two in a random permutation tree is asymptotically a third of the nodes). Figure 4.1 shows the lemniscate trees corresponding to two randomly perturbed Chebyshev polynomials of degree 30.

Table 4.1. Average value of  $N_2$  vs. degree

$n$	10	20	30	40	50	60	70	80	90	100
mean $N_2$	2.55	5.79	9.23	12.53	15.47	19.01	22.27	25.63	29.1	32.64
$n$	110	120	130	140	150	160	170	180	190	200
mean $N_2$	35.77	39.39	42.42	46.09	49.06	52.73	55.86	59.08	62.44	65.72

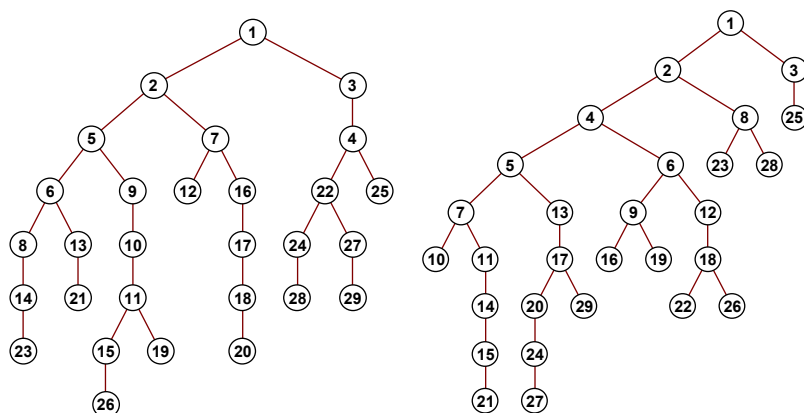


Figure 4.1. Lemniscate trees for randomly perturbed Chebyshev polynomials of degree 30.

*Remark 4.1.* — Random matrix theory gives rise to a more natural model of random polynomials that may yet exhibit a similar outcome as the perturbed Chebyshev model. Namely, consider the characteristic polynomial  $p(z) = \det(M - zI)$  of a random matrix  $M$  sampled from the so-called Jacobi ensemble [10] (with parameters chosen in order that the associated Jacobi orthogonal polynomials are Chebyshev polynomials). We expect that  $p(z)$  has a lemniscate tree with, on average, approximately one third of its nodes of outdegree two.

## BIBLIOGRAPHY

- [1] V. I. ARNOLD, “Topological classification of Morse functions and generalisations of Hilbert’s 16-th problem”, *Math. Phys. Anal. Geom.* **10** (2007), no. 3, p. 227-236.
- [2] I. BAUER & F. CATANESE, “Generic lemniscates of algebraic functions”, *Math. Ann.* **307** (1997), no. 3, p. 417-444.
- [3] F. BERGERON, P. FLAJOLET & B. SALVY, “Varieties of increasing trees”, in *Colloquium on trees in algebra and programming*, Springer, 1992, p. 24-48.
- [4] M. BÓNA, “k-Protected vertices in binary search trees”, *Adv. Appl. Math.* **53** (2014), p. 1-11.
- [5] ———, “Balanced vertices in labeled rooted trees”, <https://arxiv.org/abs/1705.09688>, 2017.
- [6] M. BÓNA & I. MEZŐ, “Limiting probabilities for vertices of a given rank in rooted trees”, <https://arxiv.org/abs/1803.05033>, 2018.
- [7] M. BÓNA & B. PITTEL, “On a random search tree: asymptotic enumeration of vertices by distance from leaves”, *Adv. Appl. Probab.* **49** (2017), no. 3, p. 850-876.
- [8] F. CATANESE, R. MIRANDA, D. ZAGIER & E. BOMBIERI, “Appendix to “Polynomial-lemniscates, trees and braids””, *Topology* **30** (1991), no. 4, p. 623-640.



- [9] F. CATANESE & M. PALUSZNY, “Polynomial-lemniscates, trees and braids”, *Topology* **30** (1991), no. 4, p. 623-640.
- [10] I. DUMITRIU & A. EDELMAN, “Matrix models for beta ensembles”, *J. Math. Phys.* **43** (2002), no. 11, p. 5830-5847.
- [11] H. EDELSBRUNNER & J. HARER, *Computational topology: an introduction*, American Mathematical Society, 2010.
- [12] P. FLAJOLET & R. SEDGEWICK, *Analytic combinatorics*, Cambridge University Press, 2009.
- [13] A. FROLOVA, D. KHAVINSON & A. VASIL'EV, “Polynomial lemniscates and their fingerprints: from geometry to topology”, in *Complex analysis and dynamical systems*, Trends in Mathematics, Birkhäuser/Springer, 2018, p. 103-128.
- [14] Y. V. FYODOROV, A. LERARIO & E. LUNDBERG, “On the number of connected components of random algebraic hypersurfaces”, *J. Geom. Phys.* **95** (2015), p. 1-20.
- [15] D. GAYET & J.-Y. WELSCHINGER, “Exponential rarefaction of real curves with many components”, *Publ. Math., Inst. Hautes Étud. Sci.* **113** (2011), no. 1, p. 69-96.
- [16] ———, “Lower estimates for the expected Betti numbers of random real hypersurfaces”, *J. Lond. Math. Soc.* **90** (2014), no. 1, p. 105-120.
- [17] ———, “Betti numbers of random real hypersurfaces and determinants of random symmetric matrices”, *J. Eur. Math. Soc.* **18** (2016), no. 4, p. 733-772.
- [18] B. HANIN, “Correlations and pairing between zeros and critical points of Gaussian random polynomials”, *Int. Math. Res. Not.* **2015** (2015), no. 2, p. 381-421.
- [19] ———, “Pairing of zeros and critical points for random meromorphic functions on Riemann surfaces”, *Math. Res. Lett.* **22** (2015), no. 1, p. 111-140.
- [20] ———, “Pairing of zeros and critical points for random polynomials”, *Ann. Inst. Henri Poincaré, Probab. Stat.* **53** (2017), no. 3, p. 1498-1511.
- [21] A. LERARIO & E. LUNDBERG, “Statistics on Hilbert’s 16th problem”, *Int. Math. Res. Not.* (2015), no. 12, p. 4293-4321.
- [22] ———, “On the geometry of random lemniscates”, *Proc. Lond. Math. Soc.* **113** (2016), no. 5, p. 649-673.
- [23] D. S. LUBINSKY, I. E. PRITSKER & X. XIE, “Expected number of real zeros for random linear combinations of orthogonal polynomials”, *Proc. Am. Math. Soc.* **144** (2016), no. 4, p. 1631-1642.
- [24] E. LUNDBERG & K. RAMACHANDRAN, “The arc length and topology of a random lemniscate”, *J. Lond. Math. Soc.* **96** (2017), no. 3, p. 621-641.
- [25] F. NAZAROV & M. SODIN, “On the number of nodal domains of random spherical harmonics”, *Am. J. Math.* **131** (2009), no. 5, p. 1337-1357.
- [26] ———, “Asymptotic laws for the spatial distribution and the number of connected components of zero sets of Gaussian random functions”, *J. Math. Phys. Anal. Geom.* **12** (2016), no. 3, p. 205-278.
- [27] L. I. NICOLAESCU, “Counting Morse functions on the 2-sphere”, *Compos. Math.* **144** (2008), no. 5, p. 1081-1106.
- [28] S. O’ROURKE & N. WILLIAMS, “Pairing between zeros and critical points of random polynomials with independent roots”, *Trans. Am. Math. Soc.* **371** (2019), no. 4, p. 2343-2381.
- [29] P. SARNAK & I. WIGMAN, “Topologies of Nodal Sets of Random Band-Limited Functions”, *Commun. Pure Appl. Math.* **72** (2019), no. 2, p. 275-342.
- [30] E. C. TITCHMARSH, *The theory of functions*, Oxford University Press, 1939.

Manuscrit reçu le 14 juillet 2018,  
révisé le 24 juin 2019,  
accepté le 11 juillet 2019.

Michael EPSTEIN  
Colorado State University  
Dept. of Mathematics  
Fort Collins, CO 80523 (USA)  
michael.epstein@colostate.edu

Boris HANIN  
Texas A&M University  
Dept. of Mathematics  
College Station, TX 77843 (USA)  
bhanin@math.tamu.edu

Erik LUNDBERG  
Florida Atlantic University  
Dept. of Mathematical Sciences  
Boca Raton, FL 33431 (USA)  
elundber@fau.edu