Shape measures of random increasing k-trees

Alexis Darrasse*
Sorbonne Universités
UPMC Univ. Paris 06, LIP6
F-75005, Paris, France

Hsien-Kuei Hwang
Institute of Statistical Science
Academia Sinica
Taipei 115 Taiwan

Michèle Soria*
Sorbonne Universités
UPMC Univ. Paris 06, LIP6
F-75005, Paris, France

December 19, 2015

Abstract

Random increasing k-trees represent an interesting and useful class of strongly dependent graphs that have been studied widely, including being used recently as models for complex networks. We study in this paper an informative notion called *BFS-profile* and derive, by several analytic means, asymptotic estimates for its expected value, together with the limiting distribution in certain cases; some interesting consequences predicting more precisely the shapes of random k-trees are also given. Our methods of proof rely essentially on a bijection between k-trees and ordinary trees, the resolution of linear systems, and a specially framed notion called Flajolet–Odlyzko admissibility.

Key words. Random graphs, k-trees, differential equations, limiting distribution, generating functions, BFS-profile, height, width, singularity analysis, random generation, method of moments, Flajolet–Odlyzko admissible functions.

1 Introduction

A *k-tree* is a graph reducible to a k-clique (a complete graph of k vertices) by successive removals of a vertex of degree k whose neighbors form a k-clique. This class of k-trees has been widely studied in combinatorics (for enumeration and characteristic properties [5, 50]), in graph algorithms (many NP-complete problems on graphs can be solved in polynomial time on k-trees [2]), and in many other fields where k-trees are naturally encountered (see [2]). Vertices in increasing k-trees (to be described below) turn out to be remarkably close, reflecting a strongly dependent graph structure. It is no surprise that such trees exhibit the scale-free

*This work was partially supported by ANR under the contract MAGNUM, n°2010 BLAN 0204.

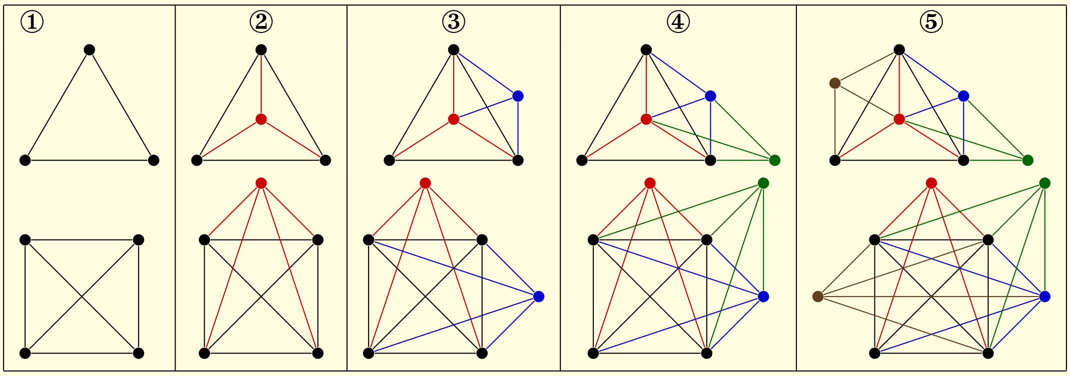


Figure 1: *The first few steps of generating a* 3*-tree (top) and a* 4*-tree (bottom).*

property [28]. Yet somewhat unexpectedly many properties of random k-trees can be dealt with by standard combinatorial, asymptotic and probabilistic tools, thus providing an important model of synergistic balance between mathematical tractability and the predictive power for practical-world complex networks.

k-trees and other recursive structures. A k-tree is not a tree (except for $k = 1$), but rather a tree-like graph that may be constructed as follows.

– *Start with an initial k-clique.*

– *Pick a k-clique in the existing k-tree, and add a new vertex joined to all k vertices in this k-clique. Repeat this step.*

At each step, exactly k new k-cliques are formed, all involving the new vertex, which may serve for later attachments (see Figure 1 for an illustration). This recursive definition facilitates the analysis and exploration of the structural properties of k-trees.

Such a simple construction process is also reminiscent of several other structures proposed in the literature such as k-DAGs [17], random circuits [3], preferential attachment [4, 9, 32], and many other models (see, for example, [8, 22, 42]). While the construction rule in each of these models is very similar, namely, linking a new vertex to k existing ones, the mechanism of choosing the existing k vertices differs from one case to another, resulting in very different topology and dynamics.

Notice that the term "k-tree" has also been used in several other areas with very different meanings; see for example [7, 23, 35, 38, 41, 46, 55].

Models of k-trees. Several variants of k-trees were proposed in the literature depending on the modeling needs and they differ mainly in the procedure of choosing a k-clique each time a new vertex is added. So k-trees can be labeled [5], unlabeled [34], increasing [57], planar [57], non-planar [5], ordered [44], or plane [43] (planar with a given embedding in the plane), etc. In particular, some families of k-trees have been employed as a model for complex networks [1, 57, 42].

For the purpose of this introduction, we distinguish between two models of random labeled non-plane k-trees; by non-plane we mean that we consider these graphs as given by a set of edges and not by its graphical representation.

Model / Properties	Simply-generated structures	Increasing structures
Combinatorial description	$\mathcal{T}_s = \mathrm{Set}(\mathcal{Z} \times \mathcal{T}_s^k)$	$\mathcal{T} = \mathrm{Set}(\mathcal{Z}^\square \times \mathcal{T}^k)$
Generating function	$T_s(z) = \exp(zT_s^k(z))$	$T'(z) = T^{k+1}(z)$
Expansion near singularity	$T_s(z) = \tau - h\sqrt{1 - z/\rho} + \cdots$	$T(z) = (1 - kz)^{-1/k}$
Mean distance of nodes	$O(\sqrt{n})$	$O(\log n)$
Degree distribution	Power law with exp. tails	Power law [28]
Root-degree distribution	Power law with exp. tails	Stable law (Theorem 12)
Expected Profile	Rayleigh limit law	Gaussian limit law (Corollary 6)

Table 1: *The contrast of some properties between random simply-generated k-trees and random increasing k-trees. Here $\mathcal{Z}$ denotes a node and $\mathcal{Z}^\square$ means a node with the smallest label.*

– The combinatorial model of *random simply-generated k-trees*, which corresponds to a uniform probability distribution on the class of labeled k-trees.

– The stochastic model of *random increasing k-trees*, where we consider the iterative generation process: at each time step, all existing k-cliques are equally likely to be selected and the new vertex is added with a label that is greater than the existing ones.

Trees and increasing trees The preceding two models are in good analogy with well-known modes of trees: the simply-generated family of trees of Meir and Moon [39], whose enumerating generating function satisfies a functional equation $f(z) = z\Phi(f(z))$, and the increasing family of trees of Bergeron, Flajolet and Salvy [6], whose generating function satisfies a differential equation $f'(z) = \Phi(f(z))$.

Very different stochastic behaviors have been observed for these families of trees: random trees belonging to the simply-generated family are often more slanted in shape, a typical description being the square-root order for the distance between two randomly chosen nodes; this is in sharp contrast with the logarithmic order for the distance of two randomly chosen nodes in random increasing trees; see for example [6, 18, 27, 37, 39, 53, 47].

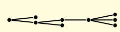
In this paper we study the class of random increasing k-trees, which is closer in behavior to random trees with logarithmic height.

Increasing k-trees versus simple k-trees. While similar in structure to trees, the analytic problems on random k-trees are however more involved because instead of a single scalar equation (either functional, algebraic, or differential), we now have a system of equations.

Table 1 presents a comparison of the two models. The class $\mathcal{T}_s$ corresponds to simply-generated k-trees and the class $\mathcal{T}$ to increasing k-trees. The results concerning simple k-trees are given in [15, 16], and those concerning increasing k-trees are derived in this paper (except for the power-law distribution [28]).

The expected distance between two randomly chosen vertices or *average path length* is one of the most important shape parameters in modeling complex networks as it indicates roughly how efficiently the information can be transmitted through the network. It is of square-root order in the simply-generated model, but of logarithmic order in the increasing model.

Another equally important parameter is the limit as $n \to \infty$ of the degree distribution of a random vertex: in the simply-generated model this is a power law with exponential tails of the form $d^{-3/2}\rho_k^d$, d denoting the degree, in contrast to a power-law of the form $d^{-1-k/(k-1)}$ [28] in the increasing model. As regards the degree of the root-vertex[1], in the simply-generated model its asymptotic distribution remains the same as that of any other vertex, but in the increasing model, the root-degree distribution is different, with an asymptotic stable law (which is Rayleigh in the case $k = 2$); see Theorem 12.

BFS-profile. Our main concern in this paper is the *BFS-profile* in increasing k-trees. Recall that the profile of a usual tree is the sequence of numbers, each enumerating the total number of nodes at a given distance to the root. For example, the tree  can be described by the profile $\{1, 2, 2, 1, 3\}$. Profiles represent one of the richest shape measures and they convey much information, particularly regarding the description of the silhouette. On random trees, they have been extensively studied recently; see [10, 19, 21, 27, 32, 37, 45]. In k-trees, the BFS-profile corresponds to the profile of the "shortest-path tree", as defined by Proskurowski [48], which is nothing more than the result of a Breadth First Search (BFS) on the graph, starting at the root-vertex. Moreover, in the domain of complex networks, this kind of BFS tree is an important object; for example, it describes the results of the `traceroute` measuring tool [52, 54] in the study of the topology of the Internet.

We derive precise asymptotic approximations for the expected BFS-profile of random increasing k-trees, the major tools used being based on the resolution of a system of differential equations of Cauchy–Euler type (see [13]). In particular, the expected number of nodes at distance d from the root-vertex follows asymptotically a Gaussian distribution (in contrast to the Rayleigh limit distribution in the case of simply-generated k-trees). Also, for d bounded, we will derive the limit distribution of the number of nodes at distance d to the root-vertex. Note that part of these results appeared as an extended abstract in [14]. For other properties of random k-trees, see [28, 40, 44, 49].

Generation of random k-trees. We simulated huge k-tree structures so that some results in the paper can be easily visualized. We also tested some of our conjectures such as the bimodality of the variance of the BFS-profile; see for example Figure 6. The generation of random k-trees can be performed by a linear-time iterative algorithm, with an amount of storage proportional to the height of the tree, and thus logarithmic in the size of the generated tree (see Corollary 7).

Organization of the paper. We first present the definition and combinatorial specification of random increasing k-trees in Section 2, together with the enumerative generating functions, on which our analytic tools will be based. In Section 3, we present two asymptotic approximations for the expected BFS-profile, one for $d = o(\log n)$ and the other for $d \to \infty$ and $d = O(\log n)$; and we also give interesting corollaries concerning expected root degree, height and width. In Section 4 we give the limit distribution of the BFS-profile in the range when $d = O(1)$.

[1]We call *root-clique* the clique composed of the k vertices $\{1, \ldots, k\}$. The increasing nature of the k-trees guarantees that these vertices always form a clique. We call *root-vertex* the vertex with label 1.

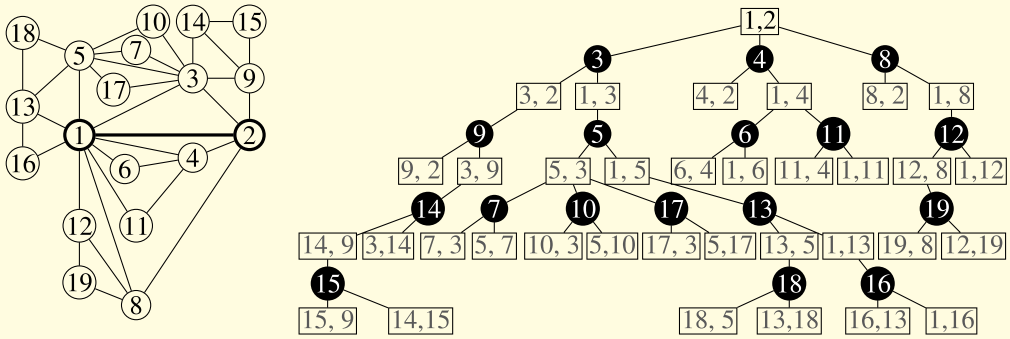


Figure 2: *A* 2*-tree (left) and its corresponding increasing tree representation (right). The* 2*-tree has* 19 *vertices and its tree representation has* 17 *black nodes. The root-vertex, labeled by* 1*, has* 10 *neighbors, including the vertex with label* 2 *which belongs to the root-clique; so it has degree* 10*, or equivalently there are* 10 *vertices at distance* 1 *to the root-vertex. All the other vertices are at distance* 2 *to the root-vertex, except for the vertex with label* 15*, which is at distance* 3*. The BFS-profile of this* 2*-tree is thus* $\{1, 10, 7, 1\}$*; its height is* 3 *and width* 10.

2 Random increasing k-trees and generating functions

Since k-trees are graphs full of cycles and cliques, the key step in our analytic-combinatorial approach is to introduce a bijection between k-trees and a suitably defined class of trees for which generating functions can be derived. This approach was successfully applied to simply-generated family of k-trees in [15], and leads to a system of algebraic equations. The bijection argument used there can be adapted *mutatis mutandis* for increasing k-trees (giving a bijection with a class of increasing trees), and it yields a system of differential equations (DEs).

2.1 Increasing k-trees and the bijection.

Let us begin by recalling the iterative construction of an increasing k-tree on $n + k$ vertices: start with a k-clique in which each vertex gets a distinct label from $\{1, \dots, k\}$; and repeatedly add n vertices as follows: at step i, select a random k-clique K_k in the existing graph and add a vertex with label $k + i$, connected by an edge to each of the k vertices of K_k. This increasing k-tree is said to be *rooted* on the initial k-clique with labels $\{1, \dots, k\}$.

In *random increasing k-trees*, we assume that all existing k-cliques are equally likely each time a new vertex is added. Notice that for $k = 1$, we get the class of plane-oriented recursive trees (see [47, 53, 36]).

The class of simply-generated k-trees (where all permutations of the labels are allowed) was studied in [15] using a combinatorial bijection with the family of trees recursively and algebraically specified by ($\mathcal{Z}$ denoting a vertex)

$$\mathcal{K}_s = \mathcal{Z}^k \times \mathcal{T}_s, \quad \text{and} \quad \mathcal{T}_s = \text{Set}\big(\mathcal{Z} \times \mathcal{T}_s^k\big).$$

Such a bijection can readily be adapted for increasing k-trees. Given a rooted increasing k-tree G with $n + k$ vertices, we can transform G into a tree structure T with black and white nodes. This structure has a white node for each k-clique of G and a black node for each vertex of G (n black nodes in total).

We start with the white node $\{1, \ldots, k\}$ corresponding to the root-clique of G, as the root of the structure T. Then we proceed recursively as follows.

- Attach to each white node labeled by $\{x_1, \ldots, x_k\}$ the unordered set of all $(k+1)$-cliques of G containing $\{x_1, \ldots, x_k\}$. Indeed, each of these $(k+1)$-cliques is represented by a black node, labeled with y which is the *only vertex* not appearing in the white-node parent.

- Each black node y is the parent of k white children, corresponding to the k other k-cliques that are formed by y and $k-1$ vertices among $\{x_1, \ldots, x_k\}$. These white children are ordered: the i-th child is the k-clique obtained by replacing x_i by y.

This procedure constructs a tree from a k-tree (see Figure 2), and, conversely, we can obtain the k-tree through a simple traversal of the tree. It is important to notice that, once given the root-clique, the labels in the white nodes are redundant since they can be calculated from the root-clique and the top-down path of their black ancestors (see [15]). The white nodes, without labels, are used to group the black children into k groups, some of which may be empty. The white nodes are not counted: only the black nodes of the tree, together with the labels in the root-clique, are needed to represent the graph. In the construction, we also see that on any path from the root to a leaf, the labels in the black nodes are in increasing order: thus the corresponding trees are *increasing trees*, respecting a monotonicity constraint on the labels.

The preceding combinatorial bijection leads to the algebraic specification ($\mathcal{Z}^{\square}$ denoting a vertex with the minimal label)

$$\mathcal{K} = \mathcal{Z}^{\square} \times \cdots \times \mathcal{Z}^{\square} \times \mathcal{T}, \quad \text{and} \quad \mathcal{T} = \text{Set}\big(\mathcal{Z}^{\square} \times \mathcal{T}^k\big). \tag{1}$$

More precisely,

- an increasing k-tree is a structure in $\mathcal{T}$, together with the sequence $\{1, \ldots, k\}$ corresponding to the labels of the root-clique. An increasing k-tree is thus completely determined by its $\mathcal{T}$-component, giving $\mathcal{K}_{n+k} \equiv \mathcal{T}_n$; the number of increasing k-trees with $n+k$ vertices is equal to the number of structures in $\mathcal{T}$ with n (black) nodes.

- The class $\mathcal{T}$ is specified by $\mathcal{T} = \text{Set}(\mathcal{Z}^{\square} \times \mathcal{T}^k)$. The product $Z^{\square} \times \mathcal{T}^k$ describes the monotonicity constraint on the labels: the boxed-product $Z^{\square} \times \mathcal{A}$ is a labeled product where the smallest label is attached to the vertex Z (see[24]). Since only the black nodes are counted, the specification describes a structure in $\mathcal{T}$ as a set of black-rooted trees, each black-rooted tree consisting of a black root with the minimal label, together with a list of k children in $\mathcal{T}$.

2.2 Generating functions.

Following the bijection, we see that the complex dependence structure of k-trees is now completely described by the class of tree-like structures specified by $\mathcal{T} = \text{Set}(\mathcal{Z}^{\square} \times \mathcal{T}^k)$.

Proposition 1. *Let $T(z) := \sum_{n\geq 0} T_n z^n/n!$ denote the exponential generating function of the number T_n of $\mathcal{T}$-structures with n vertices. Then*

$$T(z) = (1-kz)^{-1/k}, \quad \textit{and} \quad T_n = \prod_{0\leq i<n} (ik+1)$$

Recall that $T_n = K_{n+k}$: the number of increasing k-trees with $n+k$ vertices.

Proof. The box construction $\mathcal{Z}^\square \times \mathcal{A}$ translates, for generating functions, into the primitive $\int_0^z A(x)\,\mathrm{d}x$ (see[24]); thus the specification of $\mathcal{T}$ translates into the equation

$$T(z) = \exp\left(\int_0^z T^k(x)\,\mathrm{d}x\right),$$

or, equivalently, $T'(z) = T^{k+1}(z)$ with $T(0) = 1$. This implies that $T(z) = (1-kz)^{-1/k}$, and the coefficient $T_n = \prod_{0\le i<n}(ik+1)$, which is also obvious from the iterative definition, since each time a vertex is added, k new cliques appear in the random increasing k-tree. ∎

Root degree. The root degree of an increasing k-tree is the number of vertices adjacent to the root-vertex, including the $k-1$ neighbors of the root-vertex coming from the root-clique (recalling that in an increasing k-tree, the root-clique is labeled by $\{1,\ldots,k\}$, and the root-vertex is labeled by $\{1\}$).

Proposition 2. *Let $T(z,u) := \sum_{n,\ell\ge 0} T_{n,\ell} u^\ell z^n/n!$ denote the bivariate generating function of the number $T_{n,\ell}$ of increasing k-trees with $n+k$ vertices and root degree equal to $k+\ell-1$. Then*

$$T(z,u) = \left(1 - u(1-(1-kz)^{1-1/k})\right)^{-1/(k-1)}. \tag{2}$$

Proof. Consider the set of black-rooted trees at the first level of the tree, which corresponds to the $(k+1)$-cliques containing the root-clique (the three trees with roots labeled by 3, 4 and 8 in Figure 2). For each of these elements, one amongst the k subtrees is *special* since it does not contain any neighbor of the root-vertex (in the figure, for the tree with root-label 3, the special subtree is the one with root 3,2), while the $k-1$ others have the same structure as the whole tree. Marking by u the neighbors of the root-vertex in $\mathcal{T}$, we thus obtain

$$T(z,u) = \exp\left(u\int_0^z T(x)T^{k-1}(x,u)\,\mathrm{d}x\right).$$

Taking derivative with respect to z on both sides and then solving the equation, we get the closed-form expression. ∎

BFS-profile. The above bijection transforms increasing k-trees into increasing tree-structures $\mathcal{T}$; we are interested in the profile of the shortest-path trees of the k-trees, and study the corresponding parameter in trees (which is more complex than an ordinary tree-profile). Also the results provide more insight on the structure of random increasing k-trees. Roughly, we expect that all vertices in an increasing k-tree are close, one at most of logarithmic order away from the other.

Note that by construction[2], in any k-clique of the graph, if a subset S of the vertices lies at some distance b from the root-vertex, then the others lie at distance $b+1$. This means that within the "sub-tree" formed by vertices attaching to this k-clique, a vertex is at distance $b+d$ from the root-vertex if and only if it is at distance d from the set S. This leads to the consideration of the following quantities.

Consider the random variable $X_{n;d,j}$ denoting the number of nodes at distance d from a given subset S made of j vertices of the root-clique in a random increasing k-tree on $n+k$

[2] This paragraph was suggested to us by one of the referees.

vertices. Also define $X_{n;d,0}$ to be $X_{n;d-1,k}$. The bivariate generating function of the number of increasing k-trees with $n+k$ vertices and ℓ vertices at distance d from a subset of j vertices of the root-clique is $T_{d,j}(z,u) = \sum_{n\geq 0} T_n \mathbb{E}(u^{X_{n;d,j}}) z^n/n!$. By definition, $T_{d,j}(z,1) = T(z)$.

Theorem 3. *The generating functions $T_{d,j}$ satisfy the system of differential equations*

$$\frac{\partial}{\partial z} T_{d,j}(z,u) = u^{\delta_{d,1}} T_{d,j-1}^{j}(z,u) T_{d,j}^{k-j+1}(z,u), \tag{3}$$

with the initial conditions $T_{d,j}(0,u) = 1$ for $1 \leq j \leq k$, where $\delta_{a,b}$ denotes the Kronecker function, $T_{0,k}(z,u) = T(z)$ and $T_{d,0}(z,u) = T_{d-1,k}(z,u)$.

Proof. The idea of our proof is the same as in proposition 2. Let us first consider the case $d = 1$: for each black-rooted tree at the first level of the tree, $k - j$ of its white children contain all the j vertices of S and the remaining j contain all but one of them. By symmetry, the corresponding bivariate generating functions do not depend on the position of the j vertices in the root-clique. Thus the equation

$$T_{1,j}(z,u) = \exp\left(u \int_0^z T_{1,j-1}^{j}(x,u) T_{1,j}^{k-j}(x,u)\,\mathrm{d}x\right),$$

where u first marks the black nodes at the first level. For $d \geq 2$, the only difference is that the initial black nodes are not marked, thus the Kronecker function in (3). ■

Notation. For notational convenience, we normalize all z by z/k and write

$$\tilde{T}(z) := T(z/k) = (1-z)^{-1/k}.$$

Similarly, we define $\tilde{T}_{d,j}(z,u) := T_{d,j}(z/k,u)$ and have, by (3),

$$\frac{\partial}{\partial z} \tilde{T}_{d,j}(z,u) = \frac{u^{\delta_{d,1}}}{k} \tilde{T}_{d,j-1}^{j}(z,u) \tilde{T}_{d,j}^{k-j+1}(z,u), \tag{4}$$

with $\tilde{T}_{d,j}(z,1) = \tilde{T}(z)$, $\tilde{T}_{0,k}(z,u) = \tilde{T}(z)$ and $\tilde{T}_{d,0}(z,u) = \tilde{T}_{d-1,k}(z,u)$.

3 Expected BFS-profile

In this section we state our main results for the expected value of the BFS-profile $(X_{n;d,j})$, when $d = O(\log n)$. As interesting consequences, we derive (*i*) an asymptotic Gaussian approximation to the expected value of $(X_{n;d,j})$ when d is close to $(\log n)/(kH_k)$, the region where most of the vertices are concentrated, where $H_k := \sum_{1\leq j\leq k} 1/j$ denote the harmonic numbers, (*ii*) a logarithmic upper bound for the height and (*iii*) a lower bound for the width of the tree.

3.1 The expected value

We consider the expected BFS-profile, $\mathbb{E}(X_{n;d,j})$, which corresponds to the expected profile of the shortest-path trees of increasing k-trees. Observe first that

$$\mathbb{E}(X_{n;d,j}) = \frac{[z^n]\tilde{M}_{d,j}(z)}{[z^n]\tilde{T}(z)}, \quad \text{where} \quad \tilde{M}_{d,j}(z) := \left.\frac{\partial \tilde{T}_{d,j}(z,u)}{\partial u}\right|_{u=1}.$$

Here $[z^n]f(z)$ denotes the coefficient of z^n in the Taylor expansion of f.

For small d, the expected BFS-profile can be computed and estimated inductively.

Proposition 4. *For $d = O(1)$, the expected BFS-profile is asymptotically equivalent to*

$$\mathbb{E}(X_{n;d,j}) \sim \Gamma(1/k)\frac{j}{k-1} \cdot \frac{(\log n)^{d-1}}{(d-1)!} n^{1-1/k}, \tag{5}$$

for large n, fixed k, and $1 \le j \le k$.

Proof. It follows from (4) that

$$\tilde{M}'_{d,j}(z) = \frac{(k-j+1)\tilde{M}_{d,j}(z) + j\tilde{M}_{d,j-1}(z)}{k(1-z)} + \frac{\delta_{d,1}\tilde{T}(z)}{k(1-z)}. \tag{6}$$

This is a DE of standard Cauchy–Euler type whose solution is given by (see [13])

$$\tilde{M}_{d,j}(z) = \frac{(1-z)^{-(k-j+1)/k}}{k} \int_0^z (1-x)^{-(j-1)/k} \left(j\tilde{M}_{d,j-1}(x) + \delta_{d,1}\tilde{T}(x)\right) \mathrm{d}x, \tag{7}$$

since $\tilde{M}_{d,j}(0) = 0$. Then, starting from $\tilde{M}_{0,k} = 0$, we get

$$\tilde{M}_{1,1}(z) = \frac{1}{k-1}\left(\frac{1}{1-z} - \frac{1}{(1-z)^{1/k}}\right).$$

And by induction,

$$\tilde{M}_{d,j}(z) \sim \frac{j}{(k-1)(d-1)!} \cdot \frac{1}{1-z} \log^{d-1}\frac{1}{1-z},$$

for $1 \le j \le k$, $d \ge 1$ and $z \sim 1$. The asymptotics of the coefficient follows by singularity analysis (see [25, 24]). ∎

In Theorem 5 below, we show that the same asymptotic estimate holds in the larger range $d = o(\log n)$, and also derive an estimate for $d = O(\log n)$. Define the polynomial

$$P_k(z,w) := \prod_{1\le\ell\le k}\left(z - \frac{\ell}{k}\right) - c_k w = \prod_{1\le j\le k}\left(z - \lambda_j(w)\right), \tag{8}$$

where $c_k := k!/k^k$ and $\Re(\lambda_1(w)) \ge \Re(\lambda_2(w)) \ge \cdots \ge \Re(\lambda_k(w))$.

Theorem 5. *The expected BFS-profile $\mathbb{E}(X_{n;d,j})$ satisfies for $1 \le d = o(\log n)$*

$$\mathbb{E}(X_{n;d,j}) \sim \Gamma(1/k)\frac{j}{k-1} \cdot \frac{(\log n)^{d-1}}{(d-1)!} n^{1-1/k}, \tag{9}$$

uniformly in d, and for $d \to \infty$, $d = O(\log n)$,

$$\mathbb{E}(X_{n;d,j}) \sim \frac{\Gamma(1/k)h_{j,1}(\rho)\rho^{-d} n^{\lambda_1(\rho)-1/k}}{\Gamma(\lambda_1(\rho))\sqrt{2\pi(\rho\lambda_1'(\rho) + \rho^2\lambda_1''(\rho))\log n}} \tag{10}$$

where $\rho = \rho_{n,d} > 0$ solves the equation $\rho\lambda_1'(\rho) = d/\log n$, and $h_{j,1}$ is defined by

$$h_{j,1}(w) = \frac{j!w(w-1)}{(k\lambda_1(w)-1)\prod\limits_{k-j+1\le s\le k+1}(k\lambda_1(w)-s)} \cdot \frac{1}{\sum\limits_{1\le s\le k}\frac{1}{k\lambda_1(w)-s}}, \tag{11}$$

for $1 \le j \le k$.

It is far from obvious how to prove this theorem by an inductive argument of the type above, and indeed we will use an analytic approach. It is equally unclear that the denominator of (11) is nonzero when $w \ge 0$; see Section 3.3.3 for more details. Our method of proof consists of the following steps; see Figure 3. First, we consider the bivariate series $\mathscr{M}_j(z,w) := \sum_{d\ge1}\tilde{M}_{d,j}(z)w^d$, which satisfies the linear system

$$\left((1-z)\frac{\mathrm{d}}{\mathrm{d}z} - \frac{k-j+1}{k}\right)\mathscr{M}_j = \frac{j}{k}\mathscr{M}_{j-1} + \frac{w\tilde{T}}{k} \qquad (1 \le j \le k).$$

Second, this system is solved and has the solutions

$$\mathscr{M}_j(z,w) = \sum_{1\le j\le k} h_{j,m}(w)(1-z)^{-\lambda_m(w)} - \tilde{T}(z)\,\frac{w-(w-1)\delta_{k,j}}{k} \qquad (1 \le j \le k),$$

where the $h_{j,m}$'s have the same definition as $h_{j,1}$ but with all $\lambda_1(w)$ in (11) replaced by $\lambda_m(w)$. While the form of the solution is generally easy to guess, the hard part lies in the calculations of the coefficient-functions $h_{j,m}$. Third, by singularity analysis and a detailed study of the zeros, we then deduce, by the saddle-point method, the estimates given in the theorem.

Theorem 5 is proved in Section 3.3.

3.2 Corollaries

The asymptotic approximations to the mean profile entail interesting properties concerning the "shape" of random increasing k-trees, notably the distances to the root-vertex. Proofs will be given in Section 3.4.

Gaussian approximation. As a consequence of expression (10), the expected BFS-profile can be approximated by a Gaussian law when d is around $(\log n)/(kH_k)$, justifying the last item corresponding to increasing structures in Table 1. More precisely, let $H_k^{(2)} := \sum_{1\le\ell\le k} 1/\ell^2$.

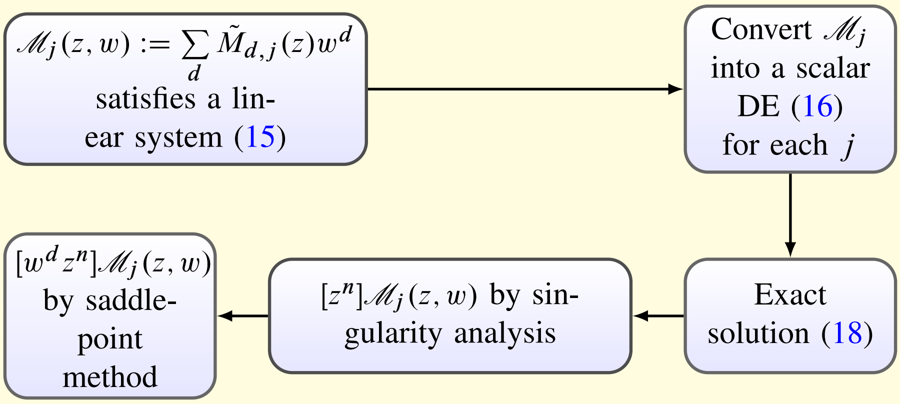


Figure 3: *Major steps used in the proof of Theorem 5.*

Corollary 6. *The expected number of nodes at distance* $d = \left\lfloor \frac{1}{kH_k}\log n + x\sigma\sqrt{\log n}\right\rfloor$ *from the root-vertex, where* $\sigma = \sqrt{H_k^{(2)}/(kH_k^3)}$*, satisfies, uniformly for* $x = o((\log n)^{1/6})$,

$$\mathbb{E}(X_{n;d,j}) \sim \frac{ne^{-x^2/2}}{\sqrt{2\pi\sigma^2\log n}}. \tag{12}$$

Expected height and width. The height of an increasing k-tree is the maximum distance to the root-vertex, or, equivalently, the number of layers in the BFS-profile plus one. On the other hand, the width is defined to be the size of the largest layer (with the maximum number of vertices) in the BFS-profile. The estimates we derived for the expected profile provide a logarithmic upper bound for the height, justifying that the mean distance of random k-trees are of logarithmic order in size, as stated in Table 1. We also get a lower bound for the width (of the shortest-path tree) of a random increasing k-tree.

Corollary 7. *Let* $\mathscr{H}_n$ *denote the height of a random increasing* k*-tree on* $n+k$ *vertices. Then*

$$\mathbb{E}(\mathscr{H}_n) \le \alpha_+\log n - \frac{\alpha_+}{2(\lambda_1(\alpha_+)-\frac{1}{k})}\log\log n + O(1), \tag{13}$$

where $\alpha_+ > 0$ *is the solution of the system of equations*

$$\begin{cases} \dfrac{1}{\alpha_+} = \displaystyle\sum_{1\le\ell\le k}\frac{1}{v-\frac{\ell}{k}}, \\ v - \dfrac{1}{k} - \alpha_+\displaystyle\sum_{1\le\ell\le k}\log\left(\frac{k}{\ell}v-1\right) = 0. \end{cases} \tag{14}$$

The following table gives the numerical values of α_+ for small values of k. For large k, one can show that $\alpha_+ \sim 1/(k\log 2)$ and $\lambda_1(\alpha_+) \sim 2$.

k	2	3	4	5	6
α_+	1.085480	0.656285	0.465190	0.358501	0.290847
k	7	8	9	10	20
α_+	0.244288	0.210365	0.184587	0.164356	0.077875

In addition to the height, we also have a lower bound for the width, which is the maximum number of nodes that appears on a single distance layer $\mathscr{W}_n := \max_d X_{n;d,j}$.

Corollary 8. *The width of a random increasing k-tree on $n+k$ vertices is bounded from below by*

$$\mathbb{E}(\mathscr{W}_n) = \mathbb{E}(\max_d X_{n,d}) \geq \max_d \mathbb{E}(X_{n,d}) \sim \frac{n}{\sqrt{2\pi\sigma^2 \log n}}.$$

We may conclude briefly from these results that *in the BFS trees of random increasing k-trees, almost all nodes are located at the levels with $d = \frac{1}{kH_k}\log n + O(\sqrt{\log n})$, each with approximately $n/\sqrt{\log n}$ nodes,* similar to most search trees of logarithmic height (see [27]). Our fine estimates also provide additional clarification of the small-world nature of random k-trees, notably $k = 3$ for which different powers of logarithmic order were claimed in [1] and in [56] for the average path-length over a deterministic 3-tree model (Apollonian networks). Strangely, while the models are different, our asymptotic result $\frac{2}{11}\log n$ is consistent with that derived in [56]; see also [37, 26] for related results. As the sequence $1/(kH_k)$ decreases moderately for increasing k and $\log n$ increases only smoothly, dubious conclusions may be drawn in practice based solely on numerical evidence.

3.3 Proof of Theorem 5

We prove in this section Theorem 5 on the asymptotic approximations to the expected BFS-profile. A diagrammatic sketch of the major steps of our proof is summarized in Figure 3.

3.3.1 Bivariate generating function and the linear system

Our approach to the asymptotics of $\mathbb{E}(X_{n;d,j})$ begins with the bivariate generating functions $\mathscr{M}_j(z,w) := \sum_{d\geq 1} \tilde{M}_{d,j}(z)w^d$ that gather information on the number of nodes at distance d from j vertices of the root-clique for every value of d.

Lemma 9. *Let θ denote the operator $(1-z)\mathbb{D}_z$. The bivariate generating functions $\mathscr{M}_j$ satisfy the linear system*

$$\left(\theta - \frac{k-j+1}{k}\right)\mathscr{M}_j = \mathscr{M}_{j-1} + \frac{w\tilde{T}}{k} \qquad (1 \leq j \leq k), \tag{15}$$

where $\mathscr{M}_0(z,w) = w\mathscr{M}_k(z,w)$.

Proof. By (6), we obtain for $j = 1, \dots, k$

$$\frac{\partial}{\partial z}\mathscr{M}_j(z,w) = \frac{k-j+1}{k(1-z)}\mathscr{M}_j(z,w) + \frac{j}{k(1-z)}\mathscr{M}_{j-1} + \frac{w\tilde{T}(z)}{k(1-z)}.$$

Thus (15) follows. ∎

The easy special case is when $k = 1$, which has the exact solution

$$\mathscr{M}_1(z,w) = (1-z)^{-w-1} - (1-z)^{-1},$$

leading to Stirling numbers of the first kind; see [24].

3.3.2 Converting the linear system (15) into scalar DEs

Instead of working directly with the linear system, we convert it into k scalar DEs of Cauchy–Euler type, with the same differential operator for all equations. For convenience, define the linear operator $\mathscr{L}_k := \prod_{1\le\ell\le k}\left(\theta-\frac{\ell}{k}\right)$. Recall that $c_k := k!/k^k$.

Proposition 10. *The bivariate generating functions $\mathscr{M}_j$ satisfy the following DEs*

$$\mathscr{L}_k[\mathscr{M}_j] = c_k w \mathscr{M}_j + \begin{cases} \frac{c_k}{k} w^2\tilde{T}, & \text{if } 1 \le j < k, \\ \frac{c_k}{k} w\tilde{T}, & \text{if } j = k, \end{cases} \tag{16}$$

where the initial conditions differ in each equation.

A particular solution is easily seen to be $-w(1-z)^{-1/k}/k$ for $\mathscr{M}_j$ for $1 \le j < k$, and $-(1-z)^{-1/k}/k$ for $\mathscr{M}_k$.

Our approach to converting the system (15) is to consider $\prod_{1\le\ell\le k}\left(\theta-\frac{\ell}{k}\right)\mathscr{M}_j$, and then simplify this by applying successively the equations in (15) one after another, so as to obtain a scalar DE for each $\mathscr{M}_j$.

Proof. We start with the proof for $\mathscr{M}_1$, the others then following by induction. By (15),

$$\begin{aligned}\mathscr{L}_k[\mathscr{M}_1] &= \prod_{1\le\ell<k}\left(\theta-\tfrac{\ell}{k}\right)(\theta-1)\mathscr{M}_1 \\ &= \prod_{1\le\ell<k}\left(\theta-\tfrac{\ell}{k}\right)\left(\frac{w}{k}\mathscr{M}_k + \frac{w\tilde{T}}{k}\right) \\ &= \frac{w}{k}\prod_{1\le\ell<k}\left(\theta-\tfrac{\ell}{k}\right)\mathscr{M}_k,\end{aligned}$$

since $\tilde{T}(z) = (1-z)^{-1/k}$ and thus $\prod_{1\le\ell<k}\left(\theta-\frac{\ell}{k}\right)\tilde{T} = \tilde{T}\prod_{1\le\ell<k}\left(\frac{1}{k}-\frac{\ell}{k}\right) = 0$. Iterating this procedure of linear operators, we finally get

$$\mathscr{L}_k[\mathscr{M}_1] = c_k w\mathscr{M}_1 + \frac{c_k}{2}w^2\tilde{T} + w^2\tilde{T}\sum_{2\le s<k}\frac{k!}{k^s(k-s+2)!}\prod_{s\le\ell<k}\left(\frac{1}{k}-\frac{\ell}{k}\right),$$

and the last two terms on the right-hand side simplify to $(k-1)!w^2\tilde{T}/k^k$. Although this approach also applies to other $\mathscr{M}_j$, it is simpler to use the following argument once we proved (16) for $j=1$. We assume first the form

$$\mathscr{L}_k[\mathscr{M}_j] = c_k w\mathscr{M}_j + C_{k,j}w^2\tilde{T}, \tag{17}$$

for $1 \le j < k$, which is already proved for $j=1$. Consider now $2 \le j < k$.

$$\mathscr{L}_k[\mathscr{M}_j] = \prod_{\ell\ne k-j+1}\left(\theta-\frac{\ell}{k}\right)\left(\theta-\frac{k-j+1}{k}\right)\mathscr{M}_j = \frac{j}{k}\prod_{\ell\ne k-j+1}\left(\theta-\frac{\ell}{k}\right)\mathscr{M}_{j-1}.$$

Multiplying the missing factor $\theta-(k-j+1)/k$, we obtain

$$\frac{j}{k}\mathscr{L}_k[\mathscr{M}_{j-1}] = \frac{j}{k}\left(c_k w\mathscr{M}_{j-1} + C_{k,j-1}w^2\tilde{T}\right),$$

which, on the other hand, also equals, by (17),

$$\left(\theta - \frac{k-j+1}{k}\right)\left(c_k w \mathscr{M}_j + C_{k,j} w^2 \tilde{T}\right) = c_k w \left(\frac{j}{k}\mathscr{M}_{j-1} + \frac{w\tilde{T}}{k}\right) - \frac{k-j}{k} C_{k,j} w^2 \tilde{T}.$$

Thus by equating these two last displays, we get the recurrence for $C_{k,j}$

$$C_{k,j} = \frac{1}{k-j}\left(\frac{k!}{k^k} - jC_{k,j-1}\right) \qquad (2 \le j < k),$$

with the initial condition $C_{k,1} = (k-1)!/k^k$. It is then easily verified that $C_{k,j} \equiv (k-1)!/k^k = c_k/k$ for $2 \le j < k$. ■

3.3.3 Solution to the Cauchy–Euler DEs (16)

To solve the Cauchy–Euler DEs (16), we observe that it can be proved that all zeros of the indicial polynomial $P_k(z, w)$ are simple for[3] $w > w_0$, where $w_0 > 0$ depends only on k; see Appendix I.

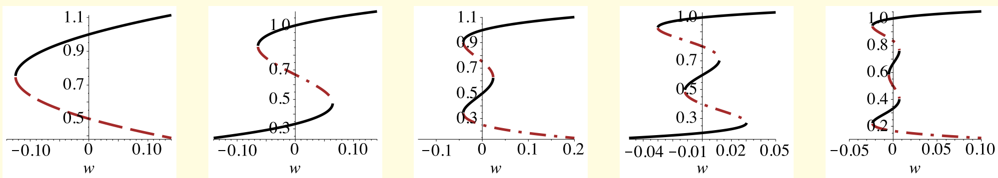


Figure 4: *A plot of the zeros $\lambda_j(w)$ as functions of w for k from 2 (leftmost) to 6 (rightmost). The top solid curve is $\lambda_1(w)$, which remains larger than 1 when $w > 0$. The other curves show that there may be double zeros (when two curves intersect) among $\{\lambda_2(w), \dots, \lambda_k(w)\}$ when w is small.*

Proposition 11. *The exact solution of (16) is given by*

$$\mathscr{M}_j(z, w) = \sum_{1 \le m \le k} h_{j,m}(w)(1-z)^{-\lambda_m(w)} - \frac{w-(w-1)\delta_{k,j}}{k}\tilde{T}(z), \tag{18}$$

where the $h_{j,m}$'s are given in (24) and (25).

Note that some of the $h_{j,m}(w)$'s are undefined at some points $w \in [-w_0, w_0]$ because their denominator becomes zero. In this case, (18) is still meaningful provided we take the limit of the right-hand side as w approaches the singularity, which results in a factor of the form $(1-z)^{-\lambda_m(w)}\log(1-z)$. Whichever the case, when $w > 0$, such a logarithmic change occurs only for lower-order terms (involving $\lambda_2(w), \dots, \lambda_k(w)$) and does not affect the dominant term (involving $\lambda_1(w)$ whose value is always larger than 1 if $w > 0$); see Figure 4 and Appendix I for details.

[3] More precisely, let z_1 denote the solution in the unit interval of the equation $\frac{1}{z} = \sum_{1 \le \ell < k} \frac{1}{\ell - z}$. Then $w_0 := \frac{k^k}{k!} \max_{\frac{1}{k} \le x \le 1} \prod_{1 \le \ell \le k} \left|1 - \frac{x}{\ell}\right| = \frac{z_1}{k} \prod_{1 \le \ell < k}\left(1 - \frac{z_1}{\ell}\right)$. For large k, $w_0 \sim e^{-1}/(k \log k)$.

In the special case when $k = 2$, the two solutions of $P_2(\theta) = 0$ are given by

$$\lambda_1(w) = \frac{3 + \sqrt{1+8w}}{4} \quad \text{and} \quad \lambda_2(w) = \frac{3 - \sqrt{1+8w}}{4},$$

and the exact solution to (16) with initial conditions zero can be easily found to be

$$\begin{aligned}
\mathscr{M}_1(z,w) &= w\frac{\lambda_1(w)(1-z)^{-\lambda_1(w)} - \lambda_2(w)(1-z)^{-\lambda_2(w)}}{\sqrt{1+8w}} - \frac{w}{2}(1-z)^{-1/2},\\
\mathscr{M}_2(z,w) &= \frac{(w-1+\lambda_1(w))(1-z)^{-\lambda_1(w)} - (w-1+\lambda_2(w))(1-z)^{-\lambda_2(w)}}{\sqrt{1+8w}} - \frac{1}{2}(1-z)^{-1/2},
\end{aligned}$$

since $\tilde{T}(z) = (1-z)^{-1/k}$; see [11, Corollary 4, p. 323]. But the expressions become more involved for higher values of k.

Proof of Proposition 11. We now prove Proposition 11 by a linear-operator approach (to characterize $h_{k,m}$) and then by induction (for other $h_{j,m}$'s). See [29, 30, 12, 13] for more information on linear-operator approach.

A linear-operator approach to deriving the coefficients $h_{j,m}$. The idea is roughly as follows. We focus on $\mathscr{M}_k$; the other cases can then be derived by (15). By (8), we define

$$\mathscr{M}_k^{[1]} := (\theta - \lambda_2(w)) \cdots (\theta - \lambda_k(w))\, \mathscr{M}_k,$$

so that the DE (16) can be written as the first-order DE $(\theta - \lambda_1(w))\mathscr{M}_k^{[1]} = c_k w T/k$, which has the solution

$$\mathscr{M}_k^{[1]}(z) = \left(\mathscr{M}_k^{[1]}(0) + \frac{c_k w}{k\lambda_1(w) - 1}\right)(1-z)^{-\lambda_1(w)} - \frac{c_k}{k\lambda_1(w)-1}(1-z)^{-1/k}. \qquad (19)$$

All coefficients are explicit except $\mathscr{M}_k^{[1]}(0)$. To derive a more explicit expression for $\mathscr{M}_k^{[1]}(0)$, we start with the identity

$$\frac{\prod_{1\le\ell\le k}\left(\theta - \frac{\ell}{k}\right) - c_k w}{\theta - \lambda} = \sum_{0\le s<k}\left[\prod_{s+2\le\ell\le k}\left(\lambda - \frac{\ell}{k}\right)\right]\left[\prod_{1\le\ell\le s}\left(\theta - \frac{\ell}{k}\right)\right], \qquad (20)$$

whenever $\lambda \in \{\lambda_1(w), \dots, \lambda_k(w)\}$, namely $P_k(\lambda, w) = 0$. To prove this, we observe that

$$\begin{aligned}
\frac{\prod_{1\le\ell\le k}\left(\theta - \frac{\ell}{k}\right)}{\theta - \lambda} &= \frac{\left[\prod_{1\le\ell<k}\left(\theta - \frac{\ell}{k}\right)\right]\left(\theta - \lambda + \lambda - \frac{k}{k}\right)}{\theta - \lambda}\\
&= \prod_{1\le\ell<k}\left(\theta - \frac{\ell}{k}\right) + (\lambda - 1)\frac{\prod_{1\le\ell<k}\left(\theta - \frac{\ell}{k}\right)}{\theta - \lambda}\\
&= \cdots\\
&= \sum_{0\le s<k}\left[\prod_{s+2\le\ell\le k}\left(\lambda - \frac{\ell}{k}\right)\right]\left[\prod_{1\le\ell\le s}\left(\theta - \frac{\ell}{k}\right)\right] + \frac{\prod_{1\le\ell\le k}\left(\lambda - \frac{\ell}{k}\right)}{\theta - \lambda}.
\end{aligned}$$

Now the last term on the right-and side will be cancelled if we add $-c_k w/(\theta-\lambda)$ to both sides because $P_k(\lambda, w) = 0$.

From (20), we have

$$\mathscr{M}_k^{[1]}(z) = \sum_{0\le s<k} \left[\prod_{s+2\le \ell\le k} \left(\lambda_1(w) - \frac{\ell}{k}\right)\right]\left[\prod_{1\le \ell\le s}\left(\theta - \frac{\ell}{k}\right)\right]\mathscr{M}_k(z).$$

This expression provides an effective means to compute the value $\mathscr{M}_k^{[1]}(0)$. To that purpose, we observe first, by induction, that

$$\prod_{1\le \ell\le s}\left(\theta - \frac{\ell}{k}\right)\mathscr{M}_k(z)\bigg|_{z=0} = \frac{(k-1)\cdots(k-s+1)}{k^s}\,w,$$

for $1 \le s \le k$. This, together with $\mathscr{M}_k(0) = 0$, gives

$$\begin{aligned}\mathscr{M}_k^{[1]}(0) &= w\sum_{1\le s<k}\left[\prod_{s+2\le \ell\le k}\left(\lambda_1(w) - \frac{\ell}{k}\right)\right]\frac{k!}{(k-s)!k^{s+1}}\\ &= c_k w \sum_{1\le s<k}\frac{\Gamma(k\lambda_1(w)-s-1)}{\Gamma(k\lambda_1(w)-k)\Gamma(k-s+1)}\\ &= \frac{c_k w(kw - k\lambda_1(w)+1)}{(k\lambda_1(w)-k-1)(k\lambda_1(w)-1)}.\end{aligned}$$

Substituting this into (19) and then iterating the linear operators, we deduce that (see, for example, [13])

$$h_{k,1}(w)\prod_{2\le \ell\le k}(\lambda_1(w)-\lambda_\ell(w)) = \frac{c_k k w(w-1)}{(k\lambda_1(w)-k-1)(k\lambda_1(w)-1)}.$$

The same procedure applies to other $h_{k,m}(w)$'s, and we obtain

$$h_{k,m}(w) = \frac{c_k k w(w-1)}{(k\lambda_m(w)-k-1)(k\lambda_m(w)-1)\prod_{\ell\ne m}(\lambda_m(w)-\lambda_\ell(w))} \qquad (1\le m\le k). \tag{21}$$

Now, again by (20),

$$\begin{aligned}\prod_{\ell\ne m}(\lambda_m(w)-\lambda_\ell(w)) &= \lim_{\theta\to\lambda_m(w)}\frac{\prod_{1\le \ell\le k}(\theta-\lambda_\ell(w))}{\theta-\lambda_m(w)}\\ &= \sum_{0\le s<k}\prod_{\ell\ne s+1}\left(\lambda_m(w)-\frac{\ell}{k}\right)\\ &= kc_k w\sum_{1\le s\le k}\frac{1}{k\lambda_m(w)-s}.\end{aligned}$$

It follows that

$$h_{k,m}(w) = \frac{w-1}{(k\lambda_m(w)-k-1)(k\lambda_m(w)-1)\sum_{1\le s\le k}\frac{1}{k\lambda_m(w)-s}} \qquad (1\le m\le k), \tag{22}$$

provided that the denominator is nonzero, which occurs when $w \sim 0$. An alternative expression is as follows.

$$h_{k,m}(w) = \frac{(w-1)\prod_{2\le s\le k}(k\lambda_m(w)-s)}{k!w(k\lambda_m(w)-k-1)\sum_{1\le s\le k}\frac{1}{k\lambda_m(w)-s}} \qquad (1 \le m \le k), \tag{23}$$

which also follows from (21). Properties of the λ_m's will be discussed in Appendix I.

Expressions for $h_{j,m}(w)$. From the linear system (15) and the expression (18), we deduce the recurrence

$$h_{j,m}(w) = \frac{j}{k\lambda_m(w)-k+j-1}\, h_{j-1,m}(w) \qquad (1 \le m, j \le k),$$

where $h_{0,m} := wh_{k,m}$. Iterating this recurrence gives

$$\begin{aligned} h_{j,m}(w) &= \frac{j!w}{\prod_{k-j+1\le s\le k}(k\lambda_m(w)-s)}\, h_{k,m}(w) \\ &= \frac{j!w(w-1)}{(k\lambda_m(w)-1)\left(\sum_{1\le s\le k}\frac{1}{k\lambda_m(w)-s}\right)\prod_{k-j+1\le s\le k+1}(k\lambda_m(w)-s)}, \end{aligned} \tag{24}$$

for $1 \le m, j \le k$, provided that the denominator is nonzero. Alternatively, we have, by (23),

$$h_{j,m}(w) = \frac{j!(w-1)\prod_{2\le s\le k-j}(k\lambda_m(w)-s)}{k!(k\lambda_m(w)-k-1)\sum_{1\le s\le k}\frac{1}{k\lambda_m(w)-s}}, \tag{25}$$

for $1 \le m \le k$ and $1 \le j < k$.

3.3.4 Properties of the zeros of P_k.

We gather in Appendix I a few properties of the zeros of $P_k(z, w)$. From Figure 5, we see that the zeros of P_k are distributed very regularly. The crucial property we need here is that $\lambda_1(w) > 1$ for $w > 0$ so that the denominators of the $h_{j,1}$'s (see (11)) are never zero when $w > 0$.

3.3.5 From $\mathscr{M}_j$ to $\tilde{M}_{d,j}$

With the expressions for the coefficient-functions $h_{j,m}$ at hand, we can now estimate the coefficient of z^n in $\tilde{M}_{d,j}$. From (18), we have, for $d \ge 1$,

$$[z^n]\tilde{M}_{d,j}(z) = [w^d z^n]\mathscr{M}_j(z, w).$$

We use singularity analysis for extracting asymptotically the coefficient of z^n in $\mathscr{M}_j(z, w)$, and then saddle-point method for the asymptotics of the coefficient of w^d in the resulting expression.

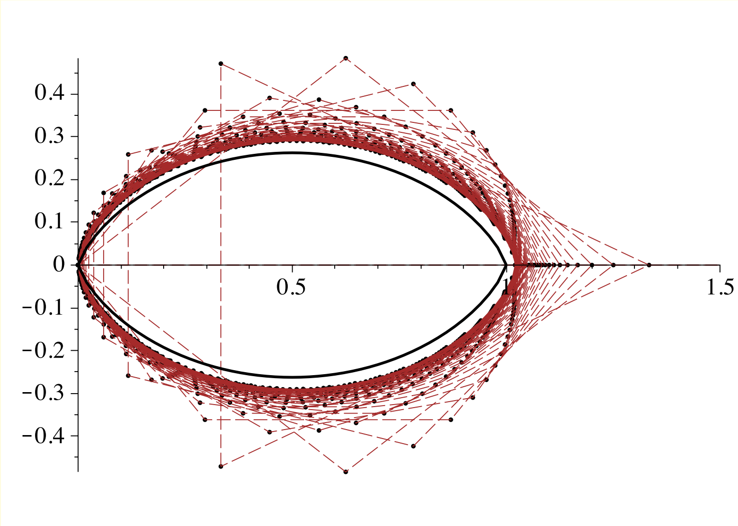


Figure 5: *Distributions of the zeros of P_k when $w = 1$ for $k = 3, \dots, 50$. The limiting curve of the zeros $|z^z(z-1)^{1-z}| = 1$ is plotted in black (innermost curve).*

Singularity analysis. From (18), we have the following identity, for $d \geq 1$

$$[z^n]\tilde{M}_{d,j}(z) = [w^d]\left(\sum_{1\leq m\leq k} h_{j,m}(w)\binom{\lambda_m(w)+n-1}{n} - \frac{w-(w-1)\delta_{k,j}}{k}\binom{\frac{1}{k}+n-1}{n}\right).$$

On the other hand, by (18) and singularity analysis, we have, for $d \geq 1$

$$[z^n]\tilde{M}_{d,j}(z) = [w^d]\left\{\frac{h_{j,1}(w)}{\Gamma(\lambda_1(w))}\, n^{\lambda_1(w)-1}\left(1+O\left(n^{-1}\right)\right)\right.$$
$$\left.+\, O\left(n^{\Re(\lambda_2(w))-1}\log n + \delta_{d,1}n^{-1+1/k}\right)\right\}.$$

Saddle-point approximations. Now by the properties of the zeros of P_k, we deduce that

$$[z^n]\tilde{M}_{d,j}(z) = \frac{1}{2\pi i}\int_{\substack{|w|=\rho\\ |\arg(w)|\leq\varepsilon}} \frac{h_{j,1}(w)}{\Gamma(\lambda_1(w))}\, w^{-d-1}n^{\lambda_1(w)-1}\left(1+O\left(n^{-\varepsilon}\right)\right) \mathrm{d}w, \qquad (26)$$

where $\rho > 0$ is chosen to be the saddle-point $\rho\lambda_1'(\rho) = d/\log n$; see [12, 31] for similar details.

Assume now $d = o(\log n)$, so that the saddle-point $\rho \sim 0$. We have

$$h_{j,1}(w) = \frac{j}{k-1}\, w\left(1+O(H_k|w|)\right) \qquad (w \sim 0).$$

Consider first $w \sim 0$. By writing

$$\prod_{1\leq\ell\leq k}\left(\theta-\frac{\ell}{k}\right) = (\theta-1)\prod_{1\leq\ell<k}\left(\theta-1+\frac{\ell}{k}\right) = \frac{k!}{k^k}(\theta-1)\prod_{1\leq\ell<k}\left(1+\frac{k}{\ell}(\theta-1)\right),$$

we have

$$\lambda_1(w)-1 = \frac{w}{\prod_{1\leq\ell<k}\left(1+\frac{k}{\ell}(\lambda_1(w)-1)\right)},$$

which, by the Lagrange inversion formula (see [24, Appendix A.5]), gives the expansion as $w \sim 0$

$$\lambda_1(w) - 1 = w - kH_{k-1}w^2 + \frac{k^2}{2}\left(3H_{k-1}^2 + H_{k-1}^{(2)}\right)w^3 + \cdots .$$

From this, it follows, by (26) and a careful error analysis, that

$$[z^n]\tilde{M}_{d,j}(z) \sim \frac{j}{k-1}[w^{d-1}]n^w = \frac{j}{k-1} \cdot \frac{(\log n)^{d-1}}{(d-1)!},$$

the error term dropped being of order $O(kH_k d/\log n)$. Thus this expression holds uniformly for $d = o(\log n)$, and the corresponding expected value is given in (9). Note that if $\alpha = d/\log n$, then $(\log n)^d/d! \asymp d^{-1/2}n^{\alpha-\alpha\log\alpha}$.

For larger values of d (indeed for $d \to \infty$), we again use (26) and apply a direct saddle-point method; the result is

$$[z^n]\tilde{M}_{d,j}(z) \sim \frac{h_{j,1}(\rho)\rho^{-d}n^{\lambda_1(\rho)-1}}{\Gamma(\lambda_1(\rho))\sqrt{2\pi(\rho\lambda_1'(\rho) + \rho^2\lambda_1''(\rho))\log n}},$$

where $\rho = \rho_{n,d} > 0$ solves the equation $\rho\lambda_1'(\rho) = d/\log n$. This gives the corresponding expected value in (10).

3.4 Proofs of Corollaries 6-7 of Theorem 5

Gaussian approximation (12) **of Corollary 6.** Since $\lambda_1(1) = (k+1)/k$, we see in this case that $\rho = 1$,

$$\rho\lambda_1'(\rho) = \frac{1}{\sum_{1\le\ell\le k}\frac{1}{\lambda_1(\rho)-\frac{\ell}{k}}},$$

and $\lambda_1(\rho) - 1/k - \alpha\log\rho \sim 1$. The more precise asymptotic estimate (12) then follows from a straightforward analysis using (10) and the local expansion

$$\lambda_1(w) - \frac{k+1}{k} = \frac{1}{kH_k}(w-1) - \frac{H_k^2 - H_k^{(2)}}{2kH_k^3}(w-1)^2 + \cdots ,$$

since $\lambda_1(1) = (k+1)/k$.

The expected height. From the estimate (10) for the expected profile, we can get an upper bound for the height. The argument is standard and as follows. Let $\mathscr{H}_n$ denote the height. Then

$$\mathbb{E}(\mathscr{H}_n) \le d_0 + \sum_{d\ge d_0}\mathbb{P}\left(\mathscr{H}_n \ge d\right),$$

for a suitably chosen d_0. So we derive an upper bound for the last sum. Let $X_{n,d}$ denote the profile (for some $j = 1, \dots, k$). Then, assuming that d_0 is an integer, we have, by the inequality $\mathbb{P}\left(\mathscr{H}_n \ge d\right) \le \sum_{\ell\ge d}\mathbb{E}(X_{n,\ell})$,

$$\sum_{d\ge d_0}\mathbb{P}\left(\mathscr{H}_n \ge d\right) \le \sum_{\ell\ge d_0}(\ell - d_0 + 1)\mathbb{E}(X_{n,\ell}) = \frac{1}{\binom{\frac{1}{k}+n-1}{n}}[z^n][w^{d_0}]\frac{\mathscr{M}(z,w)}{(1-1/w)^2},$$

where $\mathscr{M} = \mathscr{M}_j$ for some j. By our saddle-point analysis, the additional factor $(1-1/w)^2$ in the denominator does not change the asymptotic order of the expected profile if we choose d large enough (so that $\rho > 1$), and we also have, by a similar (and simpler) analysis as above,

$$\sum_{d\ge d_0} \mathbb{P}\left(\mathscr{H}_n \ge d\right) = O\left((\log n)^{-1/2}\rho^{-d_0} n^{\lambda_1(\rho)-1/k}\right),$$

where $\rho\lambda_1'(\rho) = d_0/\log n$.

For the choice of d_0, we first identify the location where the exponent of n in (10), which is $\lambda_1(\rho) - 1/k - \alpha\log\rho$, will tend to zero for the first time as ρ increases. This can be obtained by solving the system of equations (14) for positive reals (α_+, v), where $v = \lambda_1(\alpha_+)$. In this way, we obtain the table given with Corollary 7. For large k, one can show that $\alpha_+ \sim 1/(k\log 2)$ and $\lambda_1(\alpha_+) \sim 2$.

Now we take $d_0 = \lfloor \alpha_+ \log n - \beta_+ \log\log n + O(1)\rfloor$, where $\beta_+ := \frac{\alpha_+}{2(\lambda_1(\alpha_+)-\frac{1}{k})}$, and then get (13). The upper bound (13) is, up to the second-order term, expected to be tight. This proves Corollary 7. ■

4 Limiting distributions

With the availability of the bivariate generating functions (3), we can proceed further and derive the limit distribution of $X_{n;d,j}$ in the range when $d = O(1)$. The case when $d \to \infty$ is expected to be more involved; it might be possible to further extend the limit result to higher values of d by the method used in [27], but the details would be very messy.

As in the case of expected BFS-profile, we first present our results, and then give the proofs at the end of the section.

4.1 Distribution of the BFS-profile

Theorem 12. *For fixed d, the normalized random variables*

$$\bar{X}_{n;d,j} := \frac{X_{n;d,j}}{n^{1-1/k}(\log n)^{d-1}/(d-1)!}$$

converges in distribution to

$$\bar{X}_{n;d,j} \xrightarrow{d} \Xi_{d,j}, \tag{27}$$

where the limit law $\Xi_{d,j}$ has the moment generating function

$$\begin{aligned}\mathbb{E}(e^{\Xi_{d,j}u}) &= \Gamma(\tfrac{1}{k})\sum_{m\ge 0}\frac{c_{d,j,m}}{m!\Gamma(m(1-1/k)+1/k)}u^m \\ &= \frac{\Gamma(\frac{1}{k})}{2\pi i}\int_{-\infty}^{(0+)} e^{\tau}\tau^{-1/k} C_{d,j}\left(\tau^{-1+1/k}u\right)d\tau,\end{aligned}$$

and $C_{d,j}(u) := 1 + \sum_{m\ge1} c_{d,j,m}u^m/m!$ satisfies the system of DEs

$$(k-1)uC_{d,j}'(u) + C_{d,j}(u) = C_{d,j}(u)^{k+1-j}C_{d,j-1}(u)^j, \tag{28}$$

for $1 \le j \le k$, *with* $C_{d,0} = C_{d-1,k}$ *and* $C_{1,0} = 1$. *Here the symbol* $\int_{-\infty}^{(0+)}$ *denotes any Hankel contour starting from* $-\infty$ *on the real axis, encircling the origin once counter-clockwise, and returning to* $-\infty$.

Theorem 12 is proved in Section 4.3. We shall indeed prove convergence of all moments, which is stronger than weak convergence; also the limit law is uniquely determined by its moment sequence. The interesting point is that the limit laws are characterized by a system of nonlinear differential equations, bearing a similar form as the starting system of DEs (4).

Exactly solvable cases for $C_{1,j}$. Only in some special cases do we obtain explicit solutions for $C_{1,j}$

$$C_{1,1}(u) = (1-u)^{-1/(k-1)} \qquad (k \ge 2),$$
$$C_{1,2}(u) = \frac{e^{u/(1-u)}}{1-u} \qquad (k = 2),$$

and, for $k \ge 3$,

$$C_{1,2}(u) = \varphi(u)^{-\frac{1}{k-2}},$$

where for $|u| < 1$

$$\begin{aligned}\varphi(u) &:= 1 - \frac{k-2}{k-1}\sum_{m\ge1}\binom{m+\frac{2}{k-1}-1}{m}\frac{u^m}{m-\frac{k-2}{k-1}}\\ &= (1-u)^{-\frac{2}{k-1}} - 2u - \frac{2u}{k-1}\int_0^1 t^{-\frac{k-2}{k-1}}\left((1-ut)^{-\frac{2}{k-1}-1}-1\right)\mathrm{d}t.\end{aligned}$$

In particular, when $k = 3$, we have $C_{1,2}(u) = \left(1 - \frac{\sqrt{u}}{2}\log\frac{1+\sqrt{u}}{1-\sqrt{u}}\right)^{-1}$. All these solutions can be easily checked. More solutions will be discussed elsewhere.

Rayleigh distribution. Note that, when $d = 1$, the result (27) can also be derived directly by the explicit expression (2). In particular, when $k = 2$, the limit law is Rayleigh, with density $te^{-t^2/4}/2$ for $t \ge 0$, since

$$\sqrt{\pi}\sum_{m\ge0}\frac{u^m}{\Gamma((m+1)/2)} = \frac{1}{2}\int_0^\infty te^{tu-t^2/4}\,\mathrm{d}t.$$

For higher values of k, the moment generating function of Ξ is expressible in terms of generalized hypergeometric functions.

Variance. As a corollary of Theorem 12, we also get an explicit asymptotic expression for the variance of of $X_{n;d,j}$, when $d = O(1)$.

Corollary 13. *The variance of* $X_{n;d,j}$ *satisfies*

$$\mathbb{V}(X_{n;d,j}) \sim \left(\frac{\Gamma(\frac{1}{k})}{\Gamma(2-\frac{1}{k})}c_{d,j,2} - \left(\frac{j\Gamma(\frac{1}{k})}{k-1}\right)^2\right)\left(n^{1-1/k}\frac{(\log n)^{d-1}}{(d-1)!}\right)^2,$$

for $d = O(1)$ *and* $1 \le j \le k$, *where* $c_{d,j,2}$ *is given below in (40).*

This Corollary is proved in Section 4.3.4.

In what follows, we first prove Theorem 12 in the simplest case when $d = j = 1$, for which $X_{n;1,1}$ is nothing but the root degree. Our method of proof is based on exact solution and the method of moments. Then we consider the case when $d = 1$ and $1 \le j \le k$ for which closed-form solutions are too messy even when available; we thus use an inductive argument for the asymptotics of moments and still apply the method of moments. The proof is then easily amended for the remaining cases $d \ge 1$ and $1 \le j \le k$.

4.2 Degree distribution of the root

Consider the special case when $d = j = 1$. We state the following result for ease of reference although it is a special case of Theorem 12 when $d = j = 1$.

Note that from (9) with $d = 1$, we see that the expected degree of the root-vertex satisfies $\mathbb{E}(X_{n,1,1}) \sim \Gamma(1/k)n^{1-1/k}/(k-1)$.

Theorem 14. *The degree $X_{n;1,1}$ of the root-vertex in a random k-tree on $n+k$ vertices satisfies*

$$\frac{X_{n;1,1}}{n^{1-1/k}} \xrightarrow{d} \Xi_{1,1}, \tag{29}$$

where the limit law $\Xi_{1,1} = \Xi_{1,1}(k)$ has the moment generating function

$$\mathbb{E}\left(e^{\Xi_{1,1}u}\right) = \frac{\Gamma(\frac{1}{k})}{\Gamma(\frac{1}{k-1})} \sum_{m\ge 0} \frac{\Gamma(m+\frac{1}{k-1})}{\Gamma(m(1-\frac{1}{k})+\frac{1}{k})m!} u^m.$$

Bivariate generating function. By definition, $\tilde{T}_{1,1}(z,u) := \sum_{n\ge 0} \binom{n+1/k-1}{n} \mathbb{E}(u^{X_{n;1,1}}) z^n$. Note that $\tilde{T}_{1,1}(z,u) = T(z/k,u)$. Thus, by (2),

$$\tilde{T}_{1,1}(z,u) = \left(1 - u\left(1 - Z^{1-1/k}\right)\right)^{-1/(k-1)}, \quad \text{with} \quad Z = 1 - z. \tag{30}$$

Corollary 15. *The probability that the root-vertex has degree m satisfies*

$$\mathbb{P}(X_{n;1,1} = m) \sim \frac{\Gamma(m+\frac{1}{k-1})}{\Gamma(m)\Gamma(\frac{1}{k-1})} \cdot \frac{k-1}{k} n^{-1}, \tag{31}$$

uniformly for $1 \le m = o(n^{1-1/k})$.

Proof. By (30), we have, for $m \ge 1$,

$$\begin{aligned}
\mathbb{P}(X_{n;1,1} = m) &= \frac{[z^n u^m]\tilde{T}_{1,1}(z,u)}{[z^n](1-z)^{-1/k}} \\
&= \frac{\Gamma(m+\frac{1}{k-1})}{m!\Gamma(\frac{1}{k-1})\binom{n+1/k-1}{n}} [z^n]\left(1-(1-z)^{1-1/k}\right)^m \\
&\sim \frac{\Gamma(m+\frac{1}{k-1})}{m!\Gamma(\frac{1}{k-1})\binom{n+1/k-1}{n}} \cdot \frac{m(k-1)}{k\Gamma(\frac{1}{k})} n^{-2+1/k},
\end{aligned}$$

uniformly for $m = o(n^{1-1/k})$. Thus (31) follows. ■

In particular, when $m \to \infty$ and $m = o(n^{1-1/k})$, we have

$$\mathbb{P}(X_{n;1,1} = m) \sim \frac{m^{1/(k-1)}}{k\Gamma(\frac{k}{k-1})} n^{-1}.$$

One can also derive more precise approximations for larger values of m by a direct application of the saddle-point method. Also the dependence on m follows asymptotically a Beta function.

Remark. Inspired by the explicit solution (30), we can express the solution to the other $\tilde{T}_{d,j}$ as follows.

$$\tilde{T}_{d,j}(z,u) = \left(1 - \frac{k-j}{k} u^{\delta_{d,1}} \int_0^z \tilde{T}^j_{d,j-1}(t,u)\,\mathrm{d}t\right)^{-1/(k-j)}.$$

Indeed, (30) also follows from this with $d = 1$ and $j = 1$.

In particular,

$$\tilde{T}_{1,2}(z,u) = \left(1 - \frac{k-2}{k} u \int_0^z \left(1 - u\left(1 - (1-t)^{1-1/k}\right)\right)^{-2/(k-1)} \mathrm{d}t\right)^{-1/(k-2)}.$$

But the expressions for higher values of d or j soon become messy and less useful.

Method of moments. The explicit expression (30) can be re-written as

$$\begin{aligned}\tilde{T}_{1,1}(z,u) &= Z^{-1/k}\left(1 - (u-1)Z^{-1+1/k}\left(1 - Z^{1-1/k}\right)\right)^{-1/(k-1)} \\ &= Z^{-1/k} \sum_{m\geq 0} \binom{\frac{1}{k-1} + m - 1}{m} (u-1)^m Z^{-m(1-1/k)} \left(1 - Z^{1-1/k}\right)^m. \qquad (32)\end{aligned}$$

Thus all factorial moments of the random variable $X_{n;1,1}$ can be directly computed, and we have

$$\begin{aligned}\mathbb{E}(X_{n;1,1}\cdots(X_{n;1,1} - m + 1)) &= \frac{\Gamma(m + \frac{1}{k-1})}{\Gamma(\frac{1}{k-1})} \cdot \frac{[z^n]Z^{-m(1-1/k)-1/k}\left(1 - Z^{1-1/k}\right)^m}{[z^n]\tilde{T}(z)} \\ &\sim \frac{\Gamma(\frac{1}{k})\Gamma(m + \frac{1}{k-1})}{\Gamma(\frac{1}{k-1})\Gamma(m(1-\frac{1}{k}) + \frac{1}{k})} n^{m(1-1/k)},\end{aligned}$$

for $m \geq 0$. This proves the convergence of all integral moments of $X_{n;1,1}/n^{1-1/k}$ to those of Ξ. It remains the justify that such a moment sequence uniquely determines a distribution. But this is easy since the moment generating function is an entire function. The theorem then follows by applying the Frechet–Shohat moment convergence theorem. ■

Remark. Alternatively, the limit law (29) can be proved by the following argument. Let $u := e^{it/n^{1-1/k}}$. Then, by singularity analysis and by choosing a suitable Hankel contour $\mathcal{H}$,

$$\begin{aligned}\mathbb{E}(e^{iX_{n;1,1}(n)t/n^{1-1/k}}) &= \frac{\Gamma(\frac{1}{k})n^{1-1/k}}{2\pi i n} \int_{\mathcal{H}} e^{\tau}\left(1 - e^{it/n^{1-1/k}} + e^{it/n^{1-1/k}}(\tau/n)^{1-1/k}\right)^{-1/(k-1)} \\ &\qquad \times \left(1 + O(|\tau|^2 n^{-1})\right)\mathrm{d}\tau \\ &\to \frac{\Gamma(\frac{1}{k})}{2\pi i} \int_{-\infty}^{(0+)} e^{\tau}\left(\tau^{1-1/k} - it\right)^{-1/(k-1)} \mathrm{d}\tau = \mathbb{E}\left(e^{\Xi_{1,1}it}\right),\end{aligned}$$

which in the case of $k = 2$ equals the characteristic function of the Rayleigh distribution. An advantage of this analytic approach is that it can be further refined and gives a convergence rate.

4.3 Limiting distribution: $d \geq 1$ and $1 \leq j \leq k$

For higher values of d and j, we do not rely on explicit solutions but start from the recurrence relation derived from (4) and use instead an inductive argument, the crucial tool we need being the singularity analysis of Flajolet and Odlyzko for which we develop a notion of admissibility.

4.3.1 Recurrence relations

Let

$$\tilde{T}_{d,j}(z,u) = \sum_{m\geq 0} \frac{\tilde{M}_{d,j,m}(z)}{m!}(u-1)^m,$$

where the m-th factorial moment is related to $\tilde{M}_{d,j,m}$ by

$$\mathbb{E}\left(X_{n;d,j}(X_{n;d,j}-1)\cdots(X_{n;d,j}-m+1)\right) = \frac{[z^n]\tilde{M}_{d,j,m}(z)}{[z^n]\tilde{T}(z)}.$$

Then by (4), we have, for $d \neq 1$

$$\frac{\tilde{M}'_{d,j,m}(z)}{m!} = \frac{1}{k} \sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m\\ 0\leq h_1,\ldots,h_{k+1-j},i_1,\ldots,i_j\leq m}} \left(\prod_{1\leq \ell\leq k+1-j} \frac{\tilde{M}_{d,j,h_\ell}(z)}{h_\ell!}\right)\left(\prod_{1\leq \ell\leq j} \frac{\tilde{M}_{d,j-1,i_\ell}(z)}{i_\ell!}\right),$$

and for $d = 1$,

$$\begin{aligned}\frac{\tilde{M}'_{d,j,m}(z)}{m!} &= \frac{1}{k} \sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m\\ 0\leq h_1,\ldots,h_{k+1-j},i_1,\ldots,i_j\leq m}} \left(\prod_{1\leq \ell\leq k+1-j} \frac{\tilde{M}_{d,j,h_\ell}(z)}{h_\ell!}\right)\left(\prod_{1\leq \ell\leq j} \frac{\tilde{M}_{d,j-1,i_\ell}(z)}{i_\ell!}\right)\\ &\quad + \frac{1}{k} \sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m-1\\ 0\leq h_1,\ldots,h_{k+1-j},i_1,\ldots,i_j< m}} \left(\prod_{1\leq \ell\leq k+1-j} \frac{\tilde{M}_{d,j,h_\ell}(z)}{h_\ell!}\right)\left(\prod_{1\leq \ell\leq j} \frac{\tilde{M}_{d,j-1,i_\ell}(z)}{i_\ell!}\right),\end{aligned} \tag{33}$$

where the second sum will be seen to be asymptotically negligible.

4.3.2 Flajolet–Odlyzko admissible functions

For notational convenience, we introduce the following notion; see [24, §VI.10] for more general and thorough discussions.

Definition. A function f analytic inside the unit circle is said to be *$\mathscr{F}\mathscr{O}$-admissible* (following Flajolet and Odlyzko [25]) if (i) it can be analytically continued into a region of the form

$$\mathscr{R} := \{z : |z| \leq 1+\varepsilon, |\arg(z-1)| \geq \delta\} \setminus \{1\} \qquad (\varepsilon > 0; 0 < \delta < \pi/2),$$

and (ii) f satisfies

$$f(z) \sim c(1-z)^{-\alpha}\left(\log\frac{1}{1-z}\right)^{\beta} \qquad (c,\alpha,\beta \in \mathbb{C}),$$

uniformly as $z \to 1$ in $\mathscr{R}$. We write simply $f \in \mathscr{F}\mathscr{O}$ or $f \in \mathscr{F}\mathscr{O}_{\alpha,\beta}$ if we want to specify the growth order.

The usefulness of such a notion resides in the asymptotic estimate

$$[z^n]f(z) = \frac{c+o(1)}{\Gamma(\alpha)}n^{\alpha-1}(\log n)^{\beta},$$

if $\alpha \neq 0,-1,\dots,$ as well as the following closure properties.

Lemma 16. *$\mathscr{F}\mathscr{O}$-admissible functions are closed under addition, subtraction, multiplication, and differentiation; they are also closed under integration ($f \mapsto \int^z f$) if either $\alpha > 1$ or $\alpha = 1, \beta > -1$.*

Proof. Straightforward and details omitted. Note that if $f \in \mathscr{F}\mathscr{O}_{1,\beta}$, then

$$\int_{z_0}^{z} f(t)\,\mathrm{d}t \sim \begin{cases} \frac{1}{\beta+1}\left(\log\frac{1}{1-z}\right)^{\beta+1}, & \text{if } \beta > -1; \\ \log\log\frac{1}{1-z}, & \text{if } \beta = -1; \\ \text{constant}, & \text{if } \beta < -1, \end{cases}$$

as $z \to 1$ in $\mathscr{R}$. ∎

While one can include the powers of more iterated logarithmic functions in our definition of $\mathscr{F}\mathscr{O}$-admissible functions, still integration may fail to be closed for some powers.

Lemma 17. *Assume that*

$$(\theta - \nu)f = g \qquad (\nu \in \mathbb{C}),$$

in the sense of formal power series, where $\theta = (1-z)\mathbb{D}_z$. If $g \in \mathscr{F}\mathscr{O}_{\alpha,\beta}$, then $f \in \mathscr{F}\mathscr{O}$ provided that either $\nu < \alpha$ or $\nu = \alpha, \beta > -1$.

Proof. The solution of f is given explicitly by

$$f(z) = f(0)(1-z)^{-\nu} + (1-z)^{-\nu}\int_0^z (1-t)^{\nu-1}g(t)\mathrm{d}t.$$

This provides not only an analytic continuation of f into some indented disk $\mathscr{R}$ (through that of g), but also the local behavior of f as $z \to 1$ in $\mathscr{R}$. ∎

4.3.3 Method of moments

Assume now $d = 1$ and $1 \le j \le k$. From (32), we see that $\tilde{M}_{1,1,m} \in \mathscr{F}\mathscr{O}$ and $\tilde{M}_{1,1,m}$ satisfies $(Z = 1-z)$

$$\tilde{M}_{1,1,m}(z) \sim \binom{m+\frac{1}{k-1}-1}{m} Z^{-m(1-1/k)-1/k} \qquad (m \ge 0),$$

as $z \to 1$ in some indented disk $\mathscr{R}$.

We now prove by induction that $\tilde{M}_{1,j,m} \in \mathscr{F}\mathscr{O}$ and

$$\tilde{M}_{1,j,m}(z) \sim c_{1,j,m} Z^{-m(1-1/k)-1/k} \qquad (m \geq 0), \tag{34}$$

for some explicitly computable $c_{1,j,m}$ that will be specified later. By (5), we have $c_{1,j,0} = 1$ and $c_{1,j,1} = j/(k-1)$. Assume now $m \geq 2$. First of all, $\tilde{M}_{1,j,m}$ satisfies, by (33), the Cauchy–Euler differential equation

$$\left(\theta - \frac{k+1-j}{k}\right) \tilde{M}_{1,j,m}(z) = Q_{1,j,m}(z),$$

where

$$\begin{aligned}\frac{\tilde{Q}'_{1,j,m}(z)}{m!} = \frac{Z}{k} &\sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m \\ 0\leq h_1,\ldots,h_{k+1-j}<m \\ 0\leq i_1,\ldots,i_j\leq m}} \left(\prod_{1\leq \ell\leq k+1-j} \frac{\tilde{M}_{1,j,h_\ell}(z)}{h_\ell!}\right)\left(\prod_{1\leq \ell \leq j} \frac{\tilde{M}_{1,j-1,i_\ell}(z)}{i_\ell!}\right) \\ &+ \frac{Z}{k} \sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m-1 \\ 0\leq h_1,\ldots,h_{k+1-j},i_1,\ldots,i_j<m}} \left(\prod_{1\leq \ell\leq k+1-j} \frac{\tilde{M}_{1,j,h_\ell}(z)}{h_\ell!}\right)\left(\prod_{1\leq \ell \leq j} \frac{\tilde{M}_{1,j-1,i_\ell}(z)}{i_\ell!}\right).\end{aligned}$$

By Lemma 16 and induction, we have $\tilde{Q}_{1,j,m} \in \mathscr{F}\mathscr{O}$ and $\tilde{Q}_{1,j,m}(z) \sim c'_{1,j,m} Z^{-m(1-1/k)-1/k}$, where

$$c'_{1,j,m} = \frac{m!}{k} \sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m \\ 0\leq h_\ell<m,\, 0\leq i_\ell\leq m}} \left(\prod_{1\leq \ell\leq k+1-j} \frac{c_{1,j,h_\ell}}{h_\ell!}\right)\left(\prod_{1\leq \ell\leq j} \frac{c_{1,j-1,i_\ell}}{i_\ell!}\right).$$

By Lemma 17, $\tilde{M}_{1,j,m} \in \mathscr{F}\mathscr{O}$. Since

$$\tilde{M}_{d,j,m}(z) = (1-z)^{-(k+1-j)/k} \int_0^z (1-t)^{(j-1)/k} \tilde{Q}_{d,j,m}(t)\,\mathrm{d}t,$$

if we choose $c_{1,j,m}$ so that they satisfy the recurrence

$$\begin{aligned}c_{1,j,m} = \frac{1}{m(1-1/k)-1+j/k} \cdot \frac{1}{k} &\sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m \\ 0\leq h_\ell<m,\, 0\leq i_\ell\leq m}} \binom{m}{h_1,\ldots,h_{k+1-j},i_1,\ldots,i_j} \\ &\times c_{1,j,h_1}\cdots c_{1,j,h_{k+1-j}} c_{1,j-1,i_1}\cdots c_{1,d,j-1,i_j},\end{aligned}$$

then (34) follows by induction. The recurrence can be rewritten, by rearranging terms, as

$$\frac{c_{d,j,m}}{m!} = \frac{1}{m(k-1)-k+j} \sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m \\ 0\leq h_\ell<m,\, 0\leq i_\ell\leq m}} \left(\prod_{1\leq \ell\leq k+1-j} \frac{c_{d,j,h_\ell}}{h_\ell!}\right)\left(\prod_{1\leq \ell\leq j} \frac{c_{d,j-1,i_\ell}}{i_\ell!}\right). \tag{35}$$

In terms of the generating function $C_{d,j}(u) := \sum_{m\ge 0} c_{d,j,m} u^m/m!$, we have the system of DEs

$$(k-1)uC'_{1,j}(u) + C_{1,j}(u) = C_{1,j}(u)^{k+1-j} C_{1,j-1}(u)^j \qquad (1 \le j \le k),$$

with $C_{1,0}(u) = 1$ and $C_{1,j}(0) = 1$. The proof of the convergence in distribution of $X_{n;1,j}/n^{1-1/k}$ to $\Xi_{1,j}$ is complete if we prove the unique determination of the moment sequence for which it suffices to prove that the functions $C_{d,j}$ have finite radius of convergence, which will be given below for general d.

The same argument used above applies *mutatis mutandis* to $X_{n;d,j}$ for $d \ge 2$ and we deduce (27). We omit the messy details.

This gives, up to the part of unique determination of the limit law (proved below), the convergence in distribution of $X_{n;d,j}$ (normalized by the order of its mean) for $d = O(1)$ and $1 \le j \le k$.

Justification of the unique determination of the limit law. It suffices, by Carleman's criterion, to prove that

$$\frac{c_{d,j,m}}{m!} \le A^m \binom{\varepsilon + m - 1}{m}, \tag{36}$$

for all $m \ge 0$ and $1 \le j \le k$, where $A > 1$ is sufficiently large and $\varepsilon > 0$ is sufficiently small. To prove this uniform upper bound, we start from the recurrence (35). By induction

$$\begin{aligned}
\frac{c_{d,j,m}}{m!} &\le \frac{A^m}{m(k-1)-k+j} \sum_{\substack{h_1+\cdots+h_{k+1-j}+i_1+\cdots+i_j=m \\ 0\le h_\ell, i_\ell < m}} \prod_{1\le \ell\le k+1-j} \binom{\varepsilon + h_\ell - 1}{h_\ell} \prod_{1\le\ell\le j} \binom{\varepsilon + i_\ell - 1}{i_\ell} \\
&\le \frac{A^m}{m(k-1)-k+j} \left([z^m](1-z)^{-(k+1)\varepsilon} - (k+1-j)\binom{\varepsilon+m-1}{m} \right) \\
&= \frac{A^m}{m(k-1)-k+j} \left(\binom{(k+1)\varepsilon + m - 1}{m} - (k+1-j)\binom{\varepsilon+m-1}{m} \right),
\end{aligned}$$

and we need the last expression to be less than or equal to $A^m\binom{\varepsilon+m-1}{m}$, or equivalently

$$\binom{(k+1)\varepsilon + m - 1}{m} \le m(k-1)\binom{\varepsilon+m-1}{m}. \tag{37}$$

We first choose A large enough so that (36) holds for $m \le m_0$; then we choose ε small enough so that (37) holds for $m > m_0$ (always possible since the left-hand side $\asymp m^{-1+(k+1)\varepsilon}$ and the right-hand side $\asymp m^\varepsilon$ for large m). This completes the proof of (36) and thus Theorem 12. ∎

4.3.4 Variance

We now examine the variance and compute more explicitly the coefficients $c_{d,j,2}$, which, by (28), satisfy the recurrence

$$c_{d,j,2} = \frac{j}{k+j-2} c_{d,j-1,2} + \frac{j(j-1+jk(k-1))}{(k-1)^2(k+j-2)}.$$

Iterating this recurrence with respect to j yields

$$c_{d,j,2} = \frac{\Gamma(k-1)\Gamma(j+1)}{\Gamma(k+j-1)} c_{d,0,2} + \frac{j(k-1+j+jk(k-1))}{k(k-1)^2}. \tag{38}$$

If $d = 1$, then $c_{1,0,2} = c_{0,k,2} = 0$, so that $c_{1,j,2} = \frac{j(k-1+j+jk(k-1))}{k(k-1)^2}$. Assume now $d \geq 2$. Since $c_{d,0,2} = c_{d-1,k,2}$, we obtain, by substituting $j = k$ into (38),

$$c_{d,0,2} = c_{d-1,k,2} = \varrho_k c_{d-1,0,2} + \frac{k^3 - (k-1)^2}{(k-1)^2}, \tag{39}$$

where

$$\varrho_k := \frac{\Gamma(k-1)\Gamma(k+1)}{\Gamma(2k-1)}.$$

By iterating the recurrence (39) and by using the relation (see (31)) $c_{2,0,2} = c_{1,k,2} = \frac{k^3-(k-1)^2}{(k-1)^2}$, we have

$$c_{d,0,2} = \frac{k^3 - (k-1)^2}{(k-1)^2} \cdot \frac{1 - \varrho_k^{d-1}}{1 - \varrho_k},$$

for $k, d \geq 2$. In particular, since $\varrho_2 = 1$, the right-hand side is equal to $7(d-1)$. Thus we obtain for $k = 2$

$$c_{d,j,2} = 7(d-1) + \frac{j(3j+1)}{2} \qquad (j = 1, 2),$$

and for $k \geq 3$

$$c_{d,j,2} = \frac{\Gamma(k-1)\Gamma(j+1)}{\Gamma(k+j-1)} \cdot \frac{k^3 - (k-1)^2}{k-1} \cdot \frac{1 - \varrho_k^{d-1}}{1 - \varrho_k} + \frac{j(k-1+j+jk(k-1))}{k(k-1)^2}, \tag{40}$$

for $1 \leq j \leq k$. Note that if we interpret $(1 - x^{d-1})/(1-x)$ as $d-1$ when $x = 1$, then (40) also holds for $k = 2$. This proves Corollary (13). ■

The variance for higher values of d is more difficult and is expected to exhibit the same bimodal behavior as that of, say random binary search trees; see [20] and Figure 6.

Acknowledgements

We thank Hua-Huai Chern and Olivier Bodini for useful comments. We are also indebted to the referees for helpful suggestions for improvement in presentation.

References

[1] J. S. Andrade, Jr., H. J. Herrmann, R. F. S. Andrade, and L. R. da Silva. Apollonian networks: Simultaneously scale-free, small world, Euclidean, space filling, and with matching graphs. *Phys. Rev. Lett.*, 94(1):018702 (4 pages), 2005.

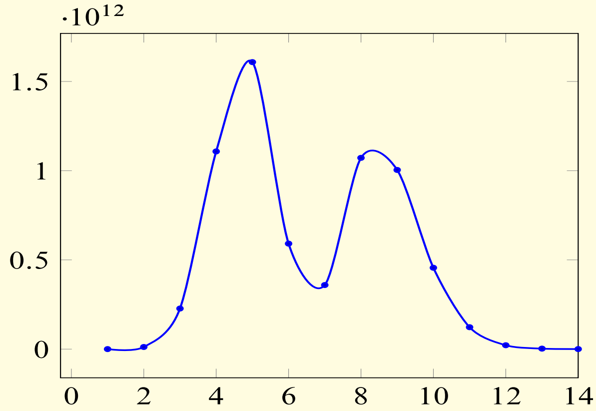


Figure 6: *The variance of the BFS-profile, calculated from* 3000 *random increasing* 2*-trees of size* $5 \cdot 10^7$.

[2] S. Arnborg, D. G. Corneil, and A. Proskurowski. Complexity of finding embeddings in a k-tree. *SIAM J. Algebraic Discrete Methods*, 8(2):277–284, 1987.

[3] S. Arya, M. J. Golin, and K. Mehlhorn. On the expected depth of random circuits. *Combin. Probab. Comput.*, 8(03):209–228, 1999.

[4] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[5] L. W. Beineke and R. E. Pippert. The number of labeled k-dimensional trees. *J. Combin. Theory*, 6(2):200–205, 1969.

[6] F. Bergeron, P. Flajolet, and B. Salvy. Varieties of increasing trees. In *CAAP '92*, pages 24–48. 1992.

[7] A. T. Berztiss. Depth-first k-trees and critical path analysis. *Acta Inform.*, 13(4):325–346, 1980.

[8] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Phys. Rep.*, 424(4-5):175–308, 2006.

[9] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 18(3):279–290, 2001.

[10] B. Chauvin, T. Klein, J.-F. Marckert, and A. Rouault. Martingales and profile of binary search trees. *Electron. J. Probab.*, 10:420–435, 2005.

[11] H.-H. Chern and H.-K. Hwang. Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms*, 19(3-4):316–358, 2001.

[12] H.-H. Chern and H.-K. Hwang. Transitional behaviors of the average cost of Quicksort with median-of-$(2t + 1)$. *Algorithmica*, 29(1-2):44–69, 2001.

[13] H.-H. Chern, H.-K. Hwang, and T.-H. Tsai. An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms. *J. Algorithms*, 44(1):177–225, 2002.

[14] A. Darrasse, H.-K. Hwang, O. Bodini, and M. Soria. The connectivity-profile of random increasing k-trees. In *Proceedings of the Seventh Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 99–106. SIAM, 2010.

[15] A. Darrasse and M. Soria. Limiting distribution for distances in k-trees. In *IWOCA 2009*, Lecture Notes in Computer Science, 5874, pages 170–182. Springer-Verlag, 2009.

[16] A. Darrasse and M. Soria. A unifying structural approach to the analysis of parameters in k-trees. *Submitted*, 2012.

[17] L. Devroye and S. Janson. Long and short paths in uniform random recursive dags. *Arkiv för Matematik*, 49(1):61–77, 2011.

[18] M. Drmota. *Random Trees*. Springer Vienna, 2009.

[19] M. Drmota and B. Gittenberger. On the profile of random trees. *Random Structures Algorithms*, 10(4):421–451, 1997.

[20] M. Drmota and H.-K. Hwang. Bimodality and phase transitions in the profile variance of random binary search trees. *SIAM J. Discrete Math.*, 19(1):19–45 (electronic), 2005.

[21] M. Drmota, S. Janson, and R. Neininger. A functional limit theorem for the profile of search trees. *Ann. Appl. Proba.*, 18(1):288–333, 2008.

[22] R. Durrett. *Random Graph Dynamics*. Cambridge University Press, 2006.

[23] M. L. Fisher. Optimal solution of vehicle routing problems using minimum k-trees. *Oper. Res.*, 42(4):626–642, 1994.

[24] F. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[25] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.

[26] A. M. Frieze and C. E. Tsourakakis. On certain properties of random apollonian networks. In A. Bonato and J. C. M. Janssen, editors, *WAW*, volume 7323 of *Lecture Notes in Computer Science*, pages 93–112. Springer, 2012.

[27] M. Fuchs, H.-K. Hwang, and R. Neininger. Profiles of random trees: Limit theorems for random recursive trees and binary search trees. *Algorithmica*, 46(3–4):367–407, 2006.

[28] Y. Gao. The degree distribution of random k-trees. *Theor. Comput. Sci.*, 410(8-10):688–695, 2009.

[29] D. H. Greene. *Labelled Formal Languages and Their Uses*. Ph.D. Thesis, Stanford University Stanford, CA, USA, 1983.

[30] P. Hennequin. *Analyse en moyenne d'algorithme, tri rapide et arbres de recherche*. Thèse, LIX, École polytechnique, 1991.

[31] H-K. Hwang. Asymptotic expansions for the Stirling numbers of the first kind. *J. Combin. Theory Ser. A*, 71(2):343–351, 1995.

[32] H.-K. Hwang. Profiles of random trees: Plane-oriented recursive trees. *Random Structures Algorithms*, 30(3):380–413, 2007.

[33] D. E. Knuth. *The Art of Computer Programming Volumes 3*. Addison-Wesley, Boston, MA, USA, second edition, 1998.

[34] G. Labelle, C. Lamathe, and P. Leroux. Labelled and unlabelled enumeration of k-gonal 2-trees. *J. Combin. Theory Ser. A*, 106(2):193–219, 2004.

[35] G. Lin. An improved approximation algorithm for multicast k-tree routing. *J. Comb. Optim.*, 9(4):349–356, 2005.

[36] H. M. Mahmoud, R. T. Smythe, and J. Szymański. On the structure of random plane-oriented recursive trees and their branches. *Random Struct. Algorithms*, 4(2):151–176, 1993.

[37] J.-F. Marckert and M. Albenque. Some families of increasing planar maps. *Electron. J. Probab.*, 13:1624–1671, 2008.

[38] C. Martinhon, A. Lucena, and N. Maculan. Stronger K-tree relaxations for the vehicle routing problem. *European J. Oper. Res.*, 158(1):56–71, 2004.

[39] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canad. J. Math.*, 30:997–1015, 1978.

[40] B. Morcrette. *Combinatoire analytique et modèles d'urnes*. Rapport, Master Parisien de Recherche en Informatique, 2010.

[41] L. Newelski and A. Rosłanowski. The ideal determined by the unsymmetric game. *Proc. Amer. Math. Soc.*, 117(3):823–831, 1993.

[42] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2):167–256, 2003.

[43] E. M. Palmer and R. C. Read. On the number of plane 2-trees. *J. London Math. Soc. (2)*, 6(4):583–592, 1973.

[44] A. Panholzer and G. Seitz. Ordered increasing k-trees: Introduction and analysis of a preferential attachment network model. In M. Drmota and B. Gittenberger, editors, *AofA'10*, pages 549–564. DMTCS, 2010.

[45] G. Park, H.-K. Hwang, P. Nicodème, and W. Szpankowski. Profiles of tries. *SIAM J. Comput.*, 38(5):1821–1880, 2009.

[46] I. Pávó. Generation of the k-trees of a graph. *Acta Cybernet.*, 1:57–68, 1971.

[47] H. Prodinger and F. J. Urbanek. On monotone functions of tree structures. *Discrete Applied Mathematics*, 5:223–239, 1983.

[48] A. Proskurowski. K-trees: representations and distances. *Congr. Numer.*, 29:785–794, 1980.

[49] A. Rényi. On the number of endpoints of a k-tree. *Studia Sci. Math. Hungar.*, 5:5–10, 1970.

[50] D. J. Rose. On simple characterizations of k-trees. *Discrete Math.*, 7(3-4):317–322, 1974.

[51] O. Roussel and M. Soria. Boltzmann sampling of ordered structures. *Electronic Notes in Discrete Mathematics*, 35:305–310, 2009.

[52] W. R. Stevens. Traceroute program. In *TCP/IP Illustrated, Volume 1: The Protocols*, chapter 8. Addison-Wesley Professional, 1994.

[53] Jerzy Szymański. On a nonuniform random recursive tree. In *Random Graphs '85 (Poznań, 1985)*, volume 144 of *North-Holland Math. Stud.*, pages 297–306. North-Holland, Amsterdam, 1987.

[54] F. Viger, B. Augustin, X. Cuvellier, C. Magnien, M. Latapy, T. Friedman, and R. Teixeira. Detection, understanding, and prevention of traceroute measurement artifacts. *Comput. Networks*, 52(5):998–1018, 2008.

[55] S. Win. On a connection between the existence of k-trees and the toughness of a graph. *Graphs Combin.*, 5(2):201–205, 1989.

[56] Z. Zhang, L. Chen, S. Zhou, L. Fang, J. Guan, and T. Zou. Analytical solution of average path length for apollonian networks. *Physical Review E*, 77(1):017102, 2008.

[57] Z. Zhang, L. Rong, and F. Comellas. High-dimensional random Apollonian networks. *Phys. A*, 364:610–618, 2006.

Appendix. I. Zero distribution of P_k

Recall that $P_k(z,w) := \prod_{1\le \ell\le k}\left(z-\frac{\ell}{k}\right) - \frac{k!}{k^k}\,w$. Most of the following properties can be proved as in [33, Ex. 6.2.4.10], [30, Annexe B], or [12].

We write for simplicity $P_k(z)$ when no ambiguity arises. Observe that

$$\frac{k^k}{k!}P_k(z,w) = \sum_{1\le j\le k+1} \mathbf{s}(k+1,j)(kz)^{j-1} - w,$$

where $\mathbf{s}(k,j)$ denotes the signless Stirling numbers of the first kind.

Z1 *If $w\in\mathbb{C}\setminus\mathbb{R}$, then all zeros of P_k are simple and not real.*

Proof. If z_0 is a zero of $P_k(z)$, then

$$\prod_{1\le\ell\le k}\left(z_0-\frac{\ell}{k}\right) = \frac{k!}{k^k}w. \tag{41}$$

Thus if w has a nonzero imaginary part, then so does z_0.

Assume now that z_0 is a zero of order higher than one, then, by (41),

$$P_k'(z_0) = \frac{k!}{k^k} w \sum_{1\le \ell \le k} \frac{1}{z_0 - \frac{\ell}{k}} = 0.$$

Since $w \in \mathbb{C}\setminus\mathbb{R}$, this implies that

$$0 = \sum_{1\le\ell\le k} \frac{1}{z_0 - \frac{\ell}{k}} = \sum_{1\le\ell\le k} \frac{\Re(z_0) - \frac{\ell}{k}}{|z_0 - \frac{\ell}{k}|^2} - i \sum_{1\le\ell\le k} \frac{\Im(z_0)}{|z_0 - \frac{\ell}{k}|^2}.$$

Thus $\Im(z_0) = 0$ and z_0 must be real, which implies in turn that w is real by (41), in contradiction to the assumption that $w \in \mathbb{C}\setminus\mathbb{R}$. Thus P_k has only simple zeros.

Z2 *For each k, there exists a $w_0 = w_0(k) > 0$ such that $P_k(z)$ has only one real zero for real w with $|w| > w_0$ if k is odd, and two or no zeros if k is even.*

Proof. If $x > 1$, then $P_k(x,0)$ is an increasing function of x. Also when k is even, if z is a zero of $P_k(z,w)$, then so is $\frac{k+1}{k} - z$. Since $P_k(j/k,0) = 0$ for $j = 1,\dots,k$, we need only to compute the quantity

$$w_0 := \frac{k^k}{k!} \max_{k^{-1}\le x\le 1} |P_k(x,0)|,$$

which will then imply the property **Z2**. We show that

$$w_0 \sim e^{-1}/(k\log k) \qquad (k\to\infty).$$

So the range of w for which P_k has more than two real zeros tends to zero with k.

Since the polynomial $P_k(z,0)$ has k distinct real zeros $\frac{1}{k},\dots,\frac{k}{k}$, we see that the equation $P_k'(z,0) = P_k'(z,w) = 0$ has $k-1$ interlacing zeros inside each of the intervals $\left(\frac{\ell}{k},\frac{\ell+1}{k}\right)$, $\ell = 1,\dots,k-1$. Denote these zeros from left to right by $\zeta_j = \zeta_{k,j}$, $j = 1,\dots,k-1$. Let also

$$w_j = w_{k,j} := \frac{k^k}{k!} P_k(\zeta_j,0) = \frac{k^k}{k!} \prod_{1\le\ell\le k}\left(\zeta_j - \frac{\ell}{k}\right) \qquad (j = 1,\dots,k-1).$$

Then $P_k(\zeta_j,w_j) = P_k'(\zeta_j,w_j) = 0$ and we have $k-1$ real roots of $P_k(z)$ of order higher than one (of order two indeed as easily checked by second derivative). Note that by symmetry, we have $w_j = (-1)^k w_{k-j}$ for $1\le j<k$. The w_j's are also (up to sign) the zeros of the resultant of $P_k(z,0)$. Thus it suffices to consider $1\le j\le k/2$. If we write

$$\zeta_j = \frac{j+z_j}{k} \qquad (0<z_j<1; 1\le j<k),$$

then

$$\begin{aligned} w_j &= \frac{k^k}{k!} \prod_{1\le\ell\le k}\left(\frac{j+z_j}{k} - \frac{\ell}{k}\right) \\ &= (-1)^{k-j}\frac{(j-1)!(k-j)!z_j}{k!}\left(\prod_{1\le\ell<j}\left(1+\frac{z_j}{\ell}\right)\right)\left(\prod_{1\le\ell\le k-j}\left(1-\frac{z_j}{\ell}\right)\right), \end{aligned}$$

where

$$\sum_{0\le \ell<j}\frac{1}{\ell+z_j}=\sum_{1\le \ell\le k-j}\frac{1}{\ell-z_j}. \tag{42}$$

If $1\le j=o(k)$, then $\log k-\log j\to\infty$, and we see that the left-hand side is asymptotic to ($H_k := \sum_{1\le j\le k} 1/j$)

$$\frac{1}{z_j}+H_j+O(z_j),$$

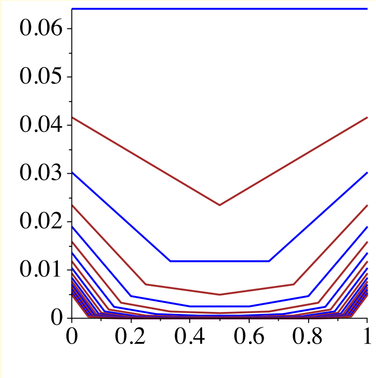
while the right-hand side is asymptotic to

$$H_k+O(jk^{-1}).$$

We thus deduce that

$$z_j\sim\frac{1}{H_k-H_j}\sim\frac{1}{\log k-\log j},$$

uniformly when $j=o(k)$. Consequently,



A plot of $|w_j|$ *for* $k=3,\ldots,10$.

$$\begin{aligned}w_j&\sim(-1)^{k-j}\frac{(j-1)!(k-j)!}{k!}z_je^{-(H_{k-j}-H_{j-1})z_j}\\&\sim(-1)^{k-j}e^{-1}\frac{(j-1)!(k-j)!}{k!\log(k/j)}.\end{aligned}$$

In particular, for finite j,

$$w_j\sim(-1)^{k-j}\frac{e^{-1}}{k^j\log k}.$$

On the other hand, when $\varepsilon k\le j\le k/2$, since the two sides of (42) are of the same order, we see that the zeros z_j remain bounded away from zero and one, and the w_j's are exponentially small in this range.

We conclude that if k is odd, then

$$w_{k-1}\le w_j\le w_1\qquad(1\le j<k; k\ge 3)$$

or $|w_j|\le w_1$, and if k is even, then

$$w_{k-1}\le w_j\le w_{k-2}\qquad(1\le j<k; k\ge 4).$$

Whichever the case,

$$w_0:=\frac{k!}{k^k}\max_{k^{-1}\le x\le 1}|P_k(x,0)|=\max_{1\le j<k}|w_j|\le|w_1|\sim e^{-1}/(k\log k).$$

Furthermore, P_k has only simple zeros for $|w|>w_0$. ■

k	w_1	w_2	w_3	w_4
2	−0.125			
3	0.06415	−0.06415		
4	−0.041666	0.0234375	−0.041666	
5	0.0302619	−0.0118224	0.0118224	−0.0302619

k	w_1	w_2	w_3	w_4
6	-0.02347346	0.00701255	-0.00488281	0.00701255
7	0.01901625	-0.00459305	0.00245213	-0.00245213
8	-0.01588792	0.00321771	-0.00139229	0.00106811
9	0.01358344	-0.00236685	0.00086060	-0.00053545

k	w_5	w_6	w_7	w_8
6	-0.02347346			
7	0.00459305	-0.01901625		
8	-0.00139229	0.00321771	-0.01588792	
9	0.00053545	-0.00086060	0.00236685	-0.01358344

Z3 *For real $w > w_0$, $\lambda_1(w)$ is an increasing function of w.*

Proof. By **Z2** and the monotonicity of $P_k(x, 0)$. ■

In particular, since $\lambda_1(0) = 1$, we see that

$$\boxed{\lambda_1(w) > 1 \quad \text{if} \quad w > 0.}$$

Z4 $\Re\lambda_j(w) = \Re\lambda_\ell(w)$ *for $j \neq \ell$ implies that $w \in \mathbb{R}$.*

Proof. By property **Z1**. ■

Note that the $\lambda_m(w)$'s are themselves analytic functions, so they vary continuously on the Riemann sphere.

Z5 *By* **Z3** *and the aperiodicity of the expected profile, $\Re\lambda_1(w) > \Re\lambda_j(w)$, for $j = 2, \ldots, k$ and $w \in \mathbb{C} \setminus \mathbb{R}$.*

Z6 $\Re\lambda_1(re^{it}) < \lambda_1(r)$, *for $r > 0$ and $0 < |t| < \pi$.*

Proof. By an argument of contradiction, we have, assuming that $\Re\lambda_1(re^{it}) \geq \lambda_1(r)$,

$$r = \prod_{1 \leq \ell \leq k} \left| \lambda_1(re^{it}) - \frac{\ell}{k} \right| > \prod_{1 \leq \ell \leq k} \left| \lambda_1(r) - \frac{\ell}{k} \right| = r,$$

because of nonzero imaginary part, and this is absurd. ■

Z7 Since $\prod_{1 \leq \ell \leq k} (\lambda_m(w) - \ell/k) = c_k w$, we see that if $w \to 0$, then one of the zeros of P_k tends to ℓ/k; by construction, $\lambda_1(w) \sim 1$ as $w \sim 0$, and for the other zeros we may assume that $\lambda_m(w) \sim (k - m + 1)/k$ as $w \sim 0$.

Z8 *As $k \to \infty$, all zeros approach the limiting curve $|z^z(z-1)^{1-z}| = 1$ when $|\arg(w)| \leq \pi - \varepsilon$, $\varepsilon > 0$; see Figure 5.*

Proof. For large k, we have, by Stirling's formula for Gamma function (with complex parameter)

$$\frac{k^k}{k!} |P_k(z, 0)| = \left| \frac{\Gamma(kz)}{k!\Gamma(kz - k)} \right| = \frac{\left|z^z(z-1)^{1-z}\right|^k \sqrt{|z-1|}}{\sqrt{2\pi k|z|}} \left(1 + O\left(\frac{1 + |z|^2}{z|z-1|k}\right)\right), \quad (43)$$

uniformly as long as $|z - 1| \gg k^{-1/2}$. This proves **Z8**. ■

Appendix. II. Random generation of increasing k-trees

We discuss briefly the large-scale simulations of random k-trees we did for the generation of the diverse figures grouped in this Appendix, which may also be useful for other purposes[4].

The generation of random k-trees and the corresponding calculation of the profiles can easily be performed by a linear-time iterative algorithm. The iterative nature of the algorithm means that a representation of the whole graph needs to be stored in memory until the generation is finished. A naive implementation of this idea then requires that we keep enough information to reconstruct the graph (for example, its tree representation); however, the maximum attainable size of the generated graphs is proportional to the amount of RAM available, which for $8 \cdot 10^7$ nodes needs about 2GB of RAM.

The space complexity of the generation can be largely reduced by only keeping track of the number of nodes at distance d to the root and with i ancestors also at distance d from the root. Such information is not only easily updated but also can be used to select a parent node when a node is generated. We thus require a storage of order $k\mathcal{H}_n$ when generating a k-tree of size n, which, according to Corollary 7 is $O(\log n)$.

Another way to reach a logarithmic space complexity would be to use a linear-complexity recursive algorithm [51], but its implementation is more involved.

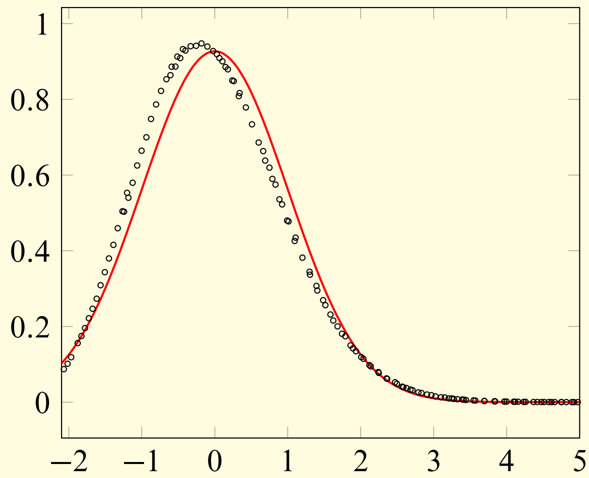


Figure 7: *A graphical rendering of Corollary 6. The simulation result of* $\mathbb{E}(X_{n;d,j})/(n/\sqrt{\log n})$ *plotted against* $x := (d - \frac{1}{2} - \frac{\log n}{kH_k})/(\sigma\sqrt{\log n})$*. The red curve represents the corresponding Gaussian density curve* $e^{-x^2/2}/\sqrt{2\pi\sigma^2}$*. The slight disagreement of the two curves is mainly due to the integer part (*d *can assume only integers).*

[4] The corresponding codes can be found in the first author's homepage.

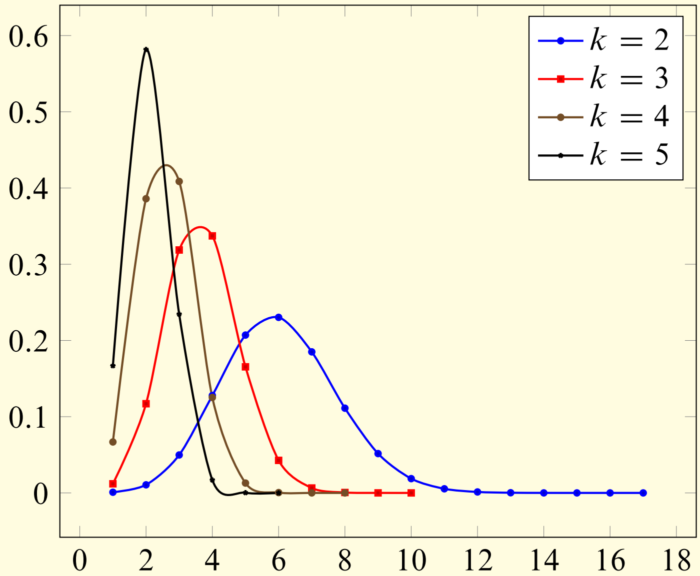


Figure 8: *The profile of a single random increasing k-tree of size $8 \cdot 10^7$ for different values of k.*

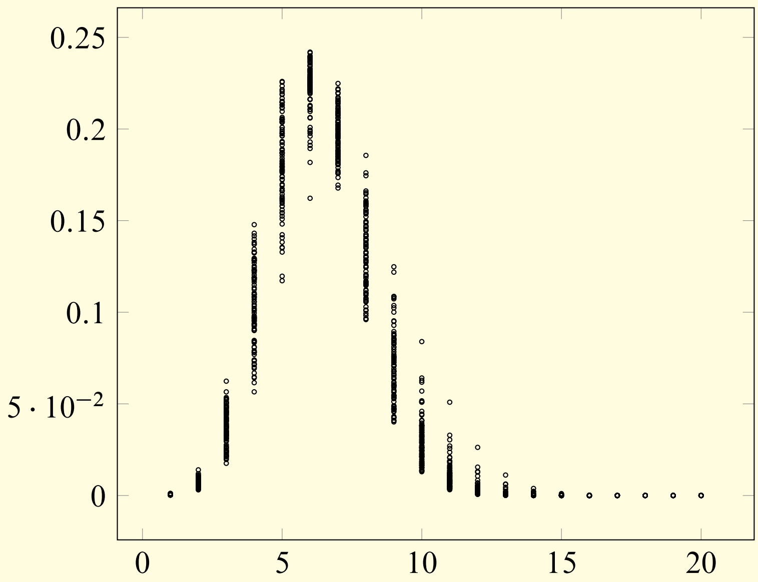


Figure 9: *The profiles of 100 random increasing 2-trees of size $5 \cdot 10^7$.*

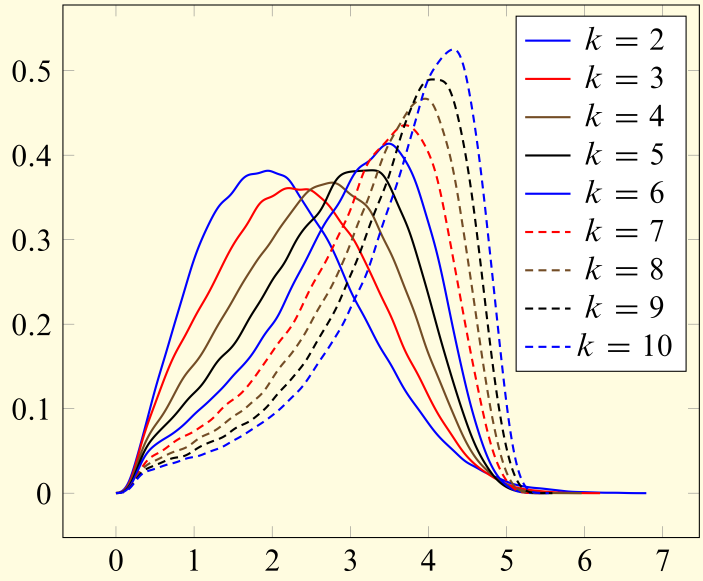


Figure 10: *Distribution of the root degree of a random increasing k-tree for k from 2 to 10.*

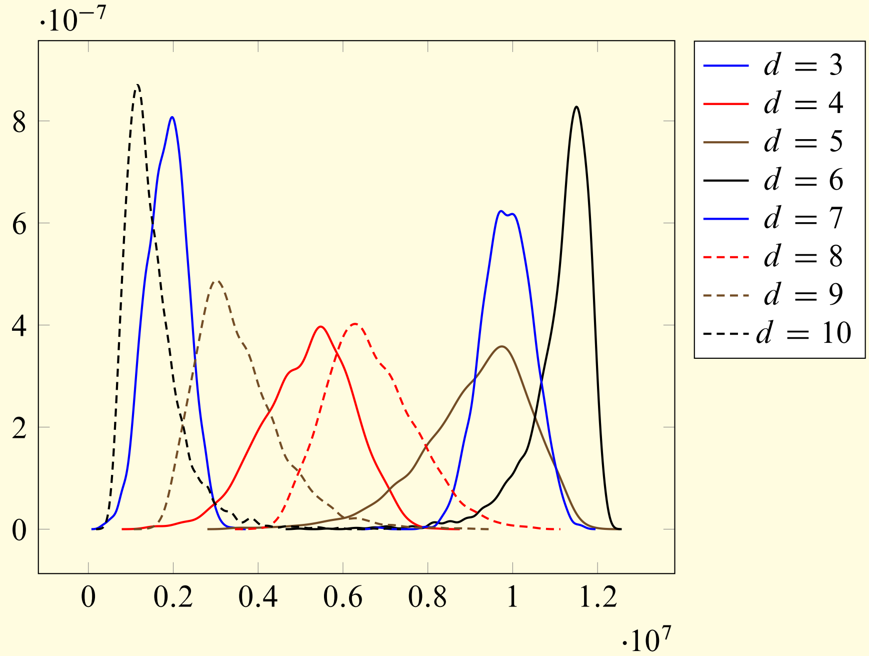


Figure 11: *Distributions of the numbers of vertices at distance d from the root, for d from 2 to 10, as calculated from 3000 random increasing 2-trees of size* $5 \cdot 10^7$.