Limiting Distribution for Distances in *k*-Trees

Alexis Darrasse and Michèle Soria*

LIP6 – UPMC Paris Universitas

Abstract. This paper examines the distances between vertices in a rooted k-tree, for a fixed k, by exhibiting a correspondence with a variety of trees that can be specified in terms of combinatorial specifications. Studying these trees via generating functions, we show a Rayleigh limiting distribution for expected distances between pairs of vertices in a random k-tree: in a k-tree on n vertices, the proportion of vertices at distance $d = x\sqrt{n}$ from a random vertex is asymptotic to $\frac{c_k^2 x}{\sqrt{n}} \exp(-\frac{c_k^2 x^2}{2})$, where $c_k = kH_k$.

1 Introduction

This work takes place within the general framework of analyzing statistical properties of combinatorial structures: we evaluate distances between vertices in a graph structure named k-tree.

In graph theory, very important research is being done on k-trees, for their characterization [15,10] and from an algorithmic viewpoint, since many NP-complete problems can be solved linearly on k-trees [1]. The class of k-trees, together with many close families, have also been extensively studied as combinatorial structures for their enumeration [2,6,9,12].

Our interest in k-trees focuses on their graph structure and the behavior of parameters such as degree or distance. This work is a generalization of our study of planar 3-trees [3], and the results presented here can be easily extended to planar k-trees.

Here we are interested in quantifying the distribution of expected distances between two random vertices in a random k-tree, for a fixed k. Our main advantage for addressing this problem is a bijection between k-trees and a class $\mathcal{K}$ which is specifiable in terms of combinatorial constructions and thus amenable to the powerful tool of generating functions which combines algebraic methods for constructing relevant power series and analytic methods for evaluating parameters of interest.

Our bijection extends works by Klawe et al. [8] and Ibarra [7] providing a one to one correspondence that exploits the recursive structure of k-cliques to transform the k-tree graph into a labeled tree structure. Moreover the parameter "k-tree distance to the root" transfers to a clearly identifiable parameter on subtrees of the tree structure, that can be precisely analysed by bivariate generating functions and leads to an asymptotic distribution that obeys a Rayleigh

* Work partially supported by ANR contract GAMMA, n°BLAN07-2_195422.

J. Fiala, J. Kratochvíl, and M. Miller (Eds.): IWOCA 2009, LNCS 5874, pp. 170–182, 2009.

law. In order to deal with distances between random pairs of vertices we use a tree "rerooting" process that is also expressible in terms of generating functions and the same result of a Rayleigh asymptotic law still holds.

Though the tree parameter that we are studying is not a profile, its behavior is similar to profiles in simple varieties of trees, where Rayleigh distributions were first shown by Meir and Moon [11], see the recent book by Drmota [4] for an extensive review.

While the class $\mathcal{K}$ that we consider here is a class of labeled trees corresponding to all k-trees on n vertices, the class of label-increasing trees of $\mathcal{K}$ would also be of interest: it would correspond, with the same bijection, to k-trees whose labeling is constrained by their recursive construction. The corresponding generating functions satisfy differential rather than algebraic equations and we show in future work that the properties are similar to those of recursive trees [4].

In section 2, we set the bijective algorithm between k-trees and class $\mathcal{K}$, and derive the enumerative generating function for k-trees. Section 3 explains the algorithm to calculate distances (in the graph) to the root of a k-tree, on the corresponding tree-structure in $\mathcal{K}$. The corresponding equations on bivariate generating functions are established by a careful analysis, marking vertices at distance d to the root and working on the cumulative generating functions to get the proportion of vertices at distance d. Finally, evaluating coefficients by means of complex analysis, we obtain the limiting distribution of distances to the root. In section 4 we extend this result to distances between a random pair of vertices in a random k-tree: we give an algorithm to "reroot" a k-tree, and show, via the equality of their generating functions, that "rerooted" k-trees are in bijection with k-trees with a pointed vertex. Using the same technique as in section 3, we finally obtain a Rayleigh limiting distribution.

2 Structures: k-Trees and Class $\mathcal{K}$

In this section we show a bijection between rooted k-trees and a simple variety of labeled trees, named $\mathcal{K}$, which is specifiable in terms of combinatorial constructions: $\mathcal{K} = \mathcal{Z}^k\mathcal{T}$, and $\mathcal{T} = \mathrm{Set}(\mathcal{Z} \times \mathcal{T}^k)$. This representation of k-trees, which highlights the cliques and their relationships, will be helpful to study the distance between vertices.

Inductive definition of k-trees. A k-tree on n vertices is a graph defined inductively as follows: the complete graph on k vertices (a k-clique) is a k-tree (with $n = k$), and if G is k-tree on $n-1$ vertices, then the graph resulting from adding a new vertex adjacent to the vertices of a k-clique of G is also a k-tree. This definition corresponds to graph theory trees when $k = 1$.

A *rooted k-tree* is a k-tree with a distinguished k-clique, together with a given permutation of the k vertices.

The $\mathcal{K}$-representation of a k-tree is a tree with black and white nodes: black nodes correspond to k+1-cliques and white nodes to k-cliques. Each black node is adjacent to the $k+1$ white nodes representing the k-cliques it contains. The size of a tree in $\mathcal{K}$ is the number of black nodes it contains plus k.

Remark 1. In this paper we shall consistently use the term *vertex* to denote "points" of the graph (k-tree), and *node* to denote "points" of the tree (in $\mathcal{K}$).

Inductive definition of $\mathcal{K}$. A tree $T \in \mathcal{K}$ of size n is either reduced to its root, a white node with k vertices (size k), or a tree $T' \in \mathcal{K}$ of size $n-1$ in which we add a new black node adjacent to a white node of T and its k white sons.

2.1 Bijection

The transformation of a rooted k-tree G into a tree $T \in \mathcal{K}$ is a two step process. First, create a tree called the completed clique-separator tree of G, whose nodes are cliques of G; moreover the root of the k-tree is the root of the corresponding tree. Second, simplify the labels of the tree by encoding most of the information in the tree structure.

The completed clique-separator tree of a k-tree G is a bipartite graph, whose black nodes are the k+1-cliques of G (which are also the maximal cliques of G), and whose white nodes are the k-cliques of G (which include the minimal separators of G). A black node is adjacent to the $k+1$ white nodes it contains. This structure is very similar to the clique-separator graph Ibarra [7] defined for the larger class of chordal graphs.

Proposition 1. *The completed clique-separator tree of* G *is a tree.*

Proof. Following the inductive definition of k-trees it is easy to see that each new vertex adds one black node connected to one existing and k new white nodes.

Proposition 2. *There is a bijection between the class of rooted* k*-trees and* $\mathcal{K}$.

Sketch of proof. For a rooted k-tree, the completed clique-separator tree is a rooted tree; its root is the root of the k-tree (a k-clique presented as an ordered list of vertices).

We see that the only information carried by a black node is the label of the vertex not included in its father. We can thus simplify the labeling of black nodes by retaining only this one vertex.

The one-to-one correspondence between k-trees and trees in $\mathcal{K}$ relies on an ordering of the sons of the black nodes. We proceed recursively from the root: consider a white node w with its ordered list of vertices $(x_1, \ldots, x_k)$; this node is adjacent to a set of black nodes and there is a natural order between the k sons of any of these black nodes. Let b be a black node: the i-th son s_i is the k-clique that does not contain x_i. The list of vertices in s_i is the same as in w except that x_i is missing and is replaced by b: $(x_1, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_k)$.

Now we only need n labels in the representation of a k-tree on n vertices; the list of vertices in each white node can be completely determined by the position of the node in the tree, the vertices in the black nodes and the list of vertices at the root (i.e.: the first white node). We do however keep the labels on the white nodes for clarity, in figure 1 for instance, they still appear, in grey color.

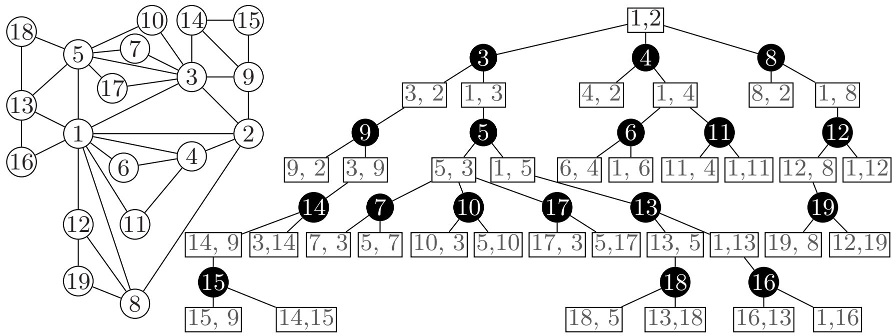

Fig. 1. A 2-tree and the corresponding tree in $\mathcal{K}$

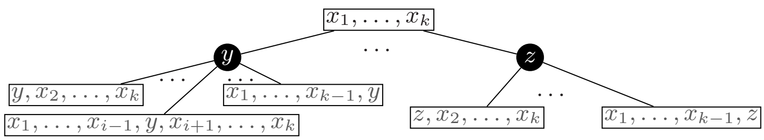

Fig. 2. The ordering of sons of black nodes

Algorithm 1. Tree representation

Input: a rooted k-tree G on n vertices
Output: a tree T, with a list of k vertices at the root

1: Create one white node for each k-clique of G
2: Create one black node for each k+1-clique of G
3: Put an edge between each black node and all k+1 k-cliques it contains
  {At this point we have the completed clique-separator tree of G}
4: Remove all vertices from each b. node except for the one not included in its father
5: Propagate the order of the vertices in white nodes starting from the root
6: Order the sons of b. nodes: the i-th does not containing the i-th vertex of the father
7: Remove all vertices from each white node {The resulting tree is in $\mathcal{K}$}

Reversing this algorithm is easy. The labels removed in steps 4 and 7 can be determined by the position of the nodes in the tree, and the list of k+1-cliques created in step 2 suffices to reconstruct the k-tree; no edge is missing: each one of them is between two vertices belonging to the same k+1-clique.

2.2 Generating Function for $\mathcal{K}$

Algorithm 1 transforms a rooted k-tree into a $\mathcal{K}$ structure, composed of a tree-structure $\mathcal{T}$ (that we call the *proper tree*), and a list of k vertices. A rooted k-tree on n vertices leads to a proper tree with $n - k$ black nodes. The proper tree is made of a white root, from which stems a set of black nodes; and each black node has a list of children which are k subtrees of the same type as the proper tree. Thus we get the following specification, with $\mathcal{E}$ denoting white nodes and $\mathcal{Z}$ denoting vertices of the original k-cliques:

$$\mathcal{K} = \mathcal{Z}^k \mathcal{T}, \qquad \mathcal{T} = \mathcal{E} \times \mathrm{Set}(\mathcal{Z} \times \mathcal{T}^k).$$

Using the symbolic method, see e.g. Flajolet and Sedgewick [5], we derive from this specification the functional equations on exponential generating functions, $K(z) = \sum_n K_n \frac{z^n}{n!}$ and $T(z) = \sum_n T_n \frac{z^n}{n!}$, where K_n (resp. T_n) is the number of trees of size n in $\mathcal{K}$ (resp. $\mathcal{T}$):

$$K(z) = z^k T(z), \qquad T(z) = \exp(zT^k(z)). \tag{1}$$

These generating functions are extensively used in the rest of this paper, especially in a bivariate form. We first use them to compute the number of k-trees on n vertices (this is another proof of a well known result [2,12, ...]).

Theorem 1. *The number of k-trees on $n+k$ vertices is $\binom{n+k}{k}(kn+1)^{n-2}$.*

Proof. Let $C_k(z) = kzT^k(z)$. Then equation (1) becomes $C_k(z) = kze^{C_k(z)}$ and $T(z) = \exp(\frac{C_k(z)}{k})$. Using the Lagrange-Bürmann theorem we thus get $[z^n]T(z) = \frac{1}{n!}(kn+1)^{n-1} = \frac{K_{n+k}}{(n+k)!}$. Since $\mathcal{K}_n$ is in bijection with the class of rooted k-trees on n vertices, we need to divide K_{n+k} by the number of possible roots, $k!(kn+1)$, to obtain the number of k-trees.

3 Distance to the Root

This section deals with distances to the first vertex of the root of a k-tree. These distances can be easily marked on the corresponding tree structure, and by studying the resulting family of bivariate generating functions, we show that the proportion of vertices at distance d from the first vertex of the root is asymptotically Rayleigh distributed.

3.1 Using the Tree to Calculate the Distance in the Graph

We use an algorithm similar to that of Proskurowski [13] to decorate each vertex with its (graph) distance to the root. Given a k-tree G and the corresponding tree T we start by assigning the distance 0 to the vertices of the root. Then, given a white node w, each of its black sons corresponds to a vertex at distance 1 plus the minimum of the distances of w's vertices.

Note that this process can be applied recursively starting from the root since each vertex in a white node w is either a part of the root or in a black node on the path from w to the root.

Algorithm 2. Distances to the first vertex of the root in $\mathcal{K}$

Input: a tree $K \in \mathcal{K}$ (with root r)
Output: an association table (vertex, distance)

1: Create an empty table A
2: **for all** vertices v in r **do**
3:    **if** v is the first vertex of r **then** Add $(v, 0)$ to A **else** Add $(v, 1)$ to A
4:    **return** $A \cup$ the result of Algorithm 3 on the proper tree of K and $(0, 1, \ldots, 1)$

Remark 2. The same process can be used to calculate the distance to any subset $s = (x_{j_1}, \ldots, x_{j_i})$ of i vertices in the root: instead of assigning distance 0 to all the vertices of the root, we set to 0 the distance of the i vertices in s and set to 1 the distance of the other $k - i$ vertices of the root. Notice that the resulting decorated tree could have been obtained as a subtree of the distance tree in the original process.

Algorithm 3. Distances in $\mathcal{T}$

Input: a tree $T \in \mathcal{T}$ and k integers $(a_i)_{i \in \{1,\ldots,k\}}$
Output: an association table (vertex, distance)

1: Let d be $1 + min(a_1, \ldots, a_k)$ and A an empty table
2: **for all** sons v of the root of T **do**
3:    Add (v, d) to A
4:    **for all** $i \in \{1, \ldots, k\}$ **do**
5:       $A \leftarrow A \cup$ recursive call on the i-th son of v and $(a_1, \ldots, a_{i-1}, d, a_{i+1}, \ldots, a_k)$
6: **return** A

Remark 3. It is clear that if we shift, by a value d, all distances of a white node w's vertices, then all the distances in the subtree under w will be consistently shifted by d.

3.2 Bivariate Generating Functions

In this section, we are interested in estimating $K_{d,n,p}$, the number of trees of size n in $\mathcal{K}$ with p vertices at distance d from the first vertex of the root. Though it is possible to directly work on the combinatorial objects, we use a generating function framework that makes presentation easier. We thus define the bivariate generating functions $K_d(z, u) = \sum_{n,p} K_{d,n,p} u^p \frac{z^n}{n!}$, with $K_d(z, 1) = K(z)$, for all d. Differentiating $K_d(z, u)$ with respect to u and setting $u = 1$, provides function

$$K'_d(z) \equiv \frac{\partial}{\partial u} K_d(z, u)\Big|_{u=1},$$

where the coefficient of $\frac{z^n}{n!}$ represents the total number of vertices at distance d from the first vertex of the root in all trees of size n in $\mathcal{K}$. So that, in a random k-tree of size n the proportion of vertices at distance d from the vertex of the root is $\frac{1}{nK_n}K'_d(z)$. The aim of this section is to give a closed form expression for $K'_d(z)$ (see proposition 3).

We mostly argument and calculate on the proper subtree T rather than on the whole structure K and the generating functions are very similar.

Lemma 1. *Let $T_d(z, u)$ be the bivariate generating function with u marking vertices at distance d from the first vertex of the root,*

$$K_1(z, u) = z^k u^{k-1} T_1(z, u), \text{ and } \forall d \geq 2,\ K_d(z, u) = z^k T_d(z, u).$$

Proof. In the case $d = 1$, each of the $k-1$ other vertices of the root are at distance one from the first vertex, since they are in the same clique. Whereas for $d \geq 2$ vertices in the root do not interfere.

Case $d = 1$. We begin with vertices at distance one to the first vertex of the root. In a second paragraph we shall be interested in vertices at distance one to a subset of i vertices in the root, in order to prepare the general study of vertices at distance d.

Lemma 2. $T_1(z, u) = \exp(zuT(z)T_1^{k-1}(z, u))$.

Proof. All sons of the root are at distance one, and for each of these black nodes, all but the first of its white children contain the first vertex of the root, so that the black nodes immediately below them are also at distance one. Hence the result for generating functions.

By differentiation we thus obtain the generating function for the total number of vertices at distance one, (also using the fact that $T_1(z, 1) = T(z)$).

Lemma 3. $T'_1(z) \equiv \frac{\partial}{\partial u} T_1(z, u)\big|_{u=1} = \frac{zT^{k+1}(z)}{1-(k-1)zT^k(z)}$.

Remark 4. If we want to count the distances to some other vertex in the root, the whole computation is the same: permuting the root's vertices brings us back to the initial problem.

The next problem is to count the number of vertices at distance one from a subset s of i vertices in the root. Extending the preceding notation ($T_1(z, u) = T_{1,1}(z, u)$), let $T_{1,i}(z, u)$ be the bivariate generating function for the number of vertices at distance one from any vertex in s.

Lemma 4. $T_{1,i}(z, u) = \exp(zuT^i_{1,i-1}(z, u)T^{k-i}_{1,i}(z, u))$.

Proof. The idea of the proof is the same as when $i = 1$: for each black node at distance one, $k-i$ of its white children contain all the i vertices of the subset and the remaining i contain all but one of them. By symmetry, the corresponding bivariate generating functions do not depend on the position of the i vertices in the root.

Differentiating leads to the generating function for the number of vertices at distance one from i vertices of the root, (notice that for all i, $T_{1,i}(z, 1) = T(z)$).

Lemma 5. $T'_{1,i}(z) \equiv \frac{\partial}{\partial u} T_{1,i}(z, u)\big|_{u=1} = \frac{zT^k(z)}{1-(k-i)zT^k(z)} \left(T(z) + iT'_{1,i-1}(z)\right)$.

This recurrence can be solved and $T'_{1,i}(z)$ is a rational function in z and $T(z)$.

General case. By remark 3, in order to calculate the number of vertices at distance two from the first vertex of the root, it suffices to find all the white nodes containing only vertices at distance one (from the first vertex of the root) and apply the previous process to count the vertices at distance one from any of the vertices of these white nodes. In terms of generating functions, this means that $T_{2,1}(z,u) = \exp(zT_{1,k}(z,u)T_{2,1}^{k-1}(z,u))$ (notice that there is no occurrence of u outside $T_{1,k}(z,u)$).

Applying the same argument recursively, we can treat the case of any distance $d \geq 2$, and obtain the following lemma.

Lemma 6. *The bivariate generating function $T_{d,i}(z,u)$, with u marking the vertices at distance d from a subset of i vertices of the root, satisfies, for $d \geq 2$:*

$$T_{d,i}(z,u) = \exp(zT_{d,i-1}^{i}(z,u)T_{d,i}^{k-i}(z,u)), \quad \textit{for} \quad i \geq 2, \quad \textit{and}$$
$$T_{d,1}(z,u) = \exp(zT_{d-1,k}(z,u)T_{d,1}^{k-1}(z,u)).$$

By differentiating we obtain:

Lemma 7. $T'_{d,i}(z) \equiv \frac{\partial}{\partial u}T_{d,i}(z,u)\big|_{u=1} = \frac{izT^k(z)}{1-(k-i)zT^k(z)}T'_{d,i-1}(z)$, *for $i \geq 1$ and $d \geq 2$, setting $T'_{d,0}(z) = T'_{d-1,k}(z)$.*

Lemma 8. $T'_{d,i}(z) = H(z)T'_{d-1,i}(z)$, *for $d \geq 2$,*
where $H(z) = k!(zT^k(z))^k \prod_{i=1}^{k-1} \frac{1}{1-izT^k(z)}$.

This recurrence is easy to expand, the only difficulty is that it does not extend to when $d = 1$. We can however calculate $T'_{2,1}(z)$ which has the form of a rational function in z and $T(z)$.

Lemma 9. $T'_{d,1}(z) \equiv \frac{\partial}{\partial u}T_{d,1}(z,u)\big|_{u=1} = H^{d-2}(z)T'_{2,1}(z)$.

Back to the whole structure $\mathcal{K}$, we have $K_d(z,u) = z^kT_{d,1}(z,u)$, so that

Proposition 3. *The exponential generating function counting the total number of vertices at distance d from the first vertex of the root in a rooted k-tree satisfies*

$$K'_d(z) \equiv \frac{\partial}{\partial u}K_d(z,u)\big|_{u=1} = H^{d-2}(z)K'_2(z),$$

with $K'_2(z) = z^kT'_{2,1}(z)$ and $H(z) = k!(zT^k(z))^k \prod_{i=1}^{k-1} \frac{1}{1-izT^k(z)}$.

3.3 Limiting Distribution

The number of vertices at distance d from the first vertex of the root in a k-tree of size n is obtained by estimating the coefficient of $\frac{z^n}{n!}$ in $K'_d(z)$ and normalizing by nK_n. We study the asymptotic of this quantity when n becomes large.

We first turn to the asymptotic estimation of coefficients of $T(z)$, $H(z)$ and $H^d(z)$, which relies on complex analysis.

Proposition 4. *Function $T(z)$ is analytic at the origin, with radius of convergence $\rho = \frac{1}{ke}$, singular value $\tau = e^{\frac{1}{k}}$, and a square-root singular expansion:*

$$T(z) = \tau - \frac{\tau}{k}\sqrt{2\left(1 - z/\rho\right)} + O\left(1 - z/\rho\right).$$

Proof. Class $\mathcal{T}$ is a simple variety of trees [11], and the result follows from the implicit function theorem. Moreover by singularity analysis $T_n \sim \frac{\tau}{k\sqrt{2\pi}}\rho^{-n}n^{-\frac{3}{2}}$.

Lemma 10. *$H(z)$ is singular in ρ, with a square-root singular expansion*

$$H(z) = 1 - kH_k\sqrt{2\left(1 - z/\rho\right)} + O\left(1 - z/\rho\right), \text{ where } H_k = \textstyle\sum_{i=1}^{k}\frac{1}{i}.$$

Proof. As seen before, $H(z) = k!(zT^k(z))^k \prod_{i=1}^{k-1} \frac{1}{1-izT^k(z)}$. For all $k \in \mathbb{N}^*$ and $i \in \{1, \ldots, k-1\}$, the term izT^k is singular in ρ and asymptotically equivalent to $i\rho\tau^k = \frac{i}{k} < 1$, so that no singularity comes from the cancellation of the denominators in the product. This product is shown to be equivalent, around the singularity ρ, to $G(z) = \frac{k^{k-1}}{(k-1)!}\prod_{i=1}^{k-1}\left(1 - \frac{i\sqrt{2(1-z/\rho)}}{k-i}\right)$. $H(z)$ is equivalent to $(zT^k(z))^k G(z)$, and the combination of square-root terms brings up a factor involving the k-th harmonic number H_k.

Evaluating the coefficient of z^n in $H^d(z)$ depends on the values of d. The region where an interesting renormalization takes place is for $d = x\sqrt{n}$, as shown in the semi-large power theorem [5, Theorem IX.16].

Proposition 5. *For $d = x\sqrt{n}$, with x in any compact of $\mathbb{R}_+^*$,*

$$[z^n]H^d(z) \sim \frac{kH_k x}{n\sqrt{2\pi}} e^{-\frac{k^2 H_k^2 x^2}{4}}.$$

Proof. This result can be obtained by using the saddle-point method or singularity analysis [5].

Back to the estimation of distances, we have $K'_d(z) = H^{d-2}(z)K'_2(z)$, where $K'_2(z)$ is a rational function of z and $T(z)$ with no pole in $[0, \rho]$, so that it contributes for a constant in the coefficient of z^n. Finally, to get the proportion of vertices at distance d, we normalize by nK_n, and obtain the following theorem.

Theorem 2. *In a k-tree on n vertices, the probability that a random vertex r is at distance $d = x\sqrt{n}$ (with x in a compact of $\mathbb{R}_+^*$) from the first vertex v of the root, satisfies, as $n \to \infty$, a local law of the Rayleigh type:*

$$\lim_{n\to\infty} \sqrt{n}\mathbb{P}(D(v,r) = \lfloor x\sqrt{n}\rfloor) = c_k^2 x e^{-\frac{(c_k x)^2}{2}}, \text{ with } c_k = k\textstyle\sum_{i=1}^{k}\frac{1}{i}.$$

In parallel to proving the limiting distribution, we made a series of measures on random k-trees generated with a purpose-built Boltzmann sampler[1], and the experimental curves perfectly fit the theoretical results, as shown in figure 3.

[1] Available at http://www-apr.lip6.fr/~darrasse/ktrees

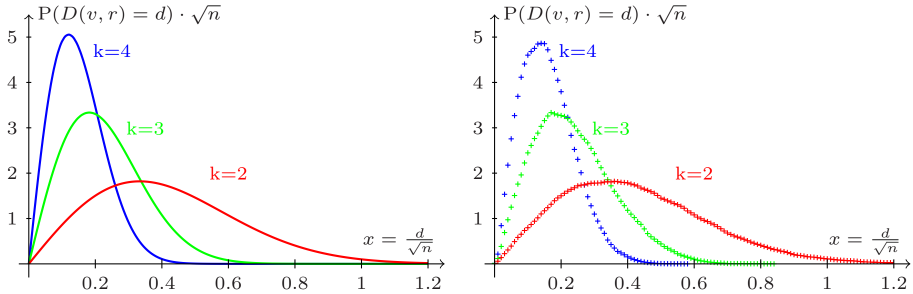

Fig. 3. Theoretical (on the left) and experimental (on the right) distributions of distances in k-trees, for $k = 2, 3$ and 4. The experimental curves come from measures on 10^3 random k-trees of size $10^4 \pm 10\%$.

4 Distances to a Random Vertex

To estimate the distance to a random vertex of a k-tree we use the same idea as Proskurowski [14]: "rerooting" the corresponding tree in order to move the vertex we are interested in to the root. In our case, we find an expression $K^\circ(z)$ of the rerooted k-trees in terms of generating functions. This is proved by exhibiting a bijection between rerooted k-trees and pointed k-trees (that is k-trees with one pointed vertex, counted by $K^\bullet(z)$).

The limiting distribution is obtained by using the same analytic tools as before and we finally prove that the distances to a random vertex exactly obey the same distribution as the previous case.

4.1 Rerooting Process

Given a distinguished vertex v in a rooted k-tree G with root r, we want to associate another rooted k-tree G' having the same underlying k-tree as G and where v is the first vertex of the new root r'. Two cases appear: either v belongs to the list of vertices of r, and exchanging v with r's first vertex suffices, or we need to find another k-clique of G to be the root of G'.

It is easier to work on the corresponding tree $T \in \mathcal{K}$. In T, we want to find a white node containing v as first vertex. There are many choices, but only one of them is always the closest to the root: the first son of v. This white node, named r', will be the root of G'. The corresponding tree T' can be either constructed using the method of section 2, or obtained directly from T by pulling on r'.

Note that this process is reversible, since the mark allows to find the old root r.

Remark 5. The mark of the old root is always contained in the first subtree of one of the sons of the new root.

Algorithm 4. Rerooting

Input: a tree $T \in \mathcal{K}$ (with root r) and a vertex v
Output: a tree in $\mathcal{K}^\circ$, with v its first root vertex

```
if v is in r then
    Put a mark on the first vertex of r
    Exchange the first vertex in r with v
    Apply steps 5 and 6 of Algorithm 1 to reorder the sons of black nodes
    return the resulting tree
else {in this case v is a black node}
    Let r' be the first son of v {r' is a white node and v its first vertex}
    Get the k-tree G corresponding to T by the inverse of Algorithm 1
    Let G' be the same graph as G, with a new root: r'
    Apply Algorithm 1 to G' to obtain a new tree T'
    return T' with a mark on the white node corresponding to r
```

Theorem 3. *The class of rerooted k-trees is counted by the generating function $K^\circ(z) = kz^kT(z) + z^{k+1}T^k(z)T^\circ(z)$, where $T^\circ(z)$ counts the trees in $\mathcal{T}$ with a mark on a white node.*

Proof. The two terms of the sum correspond to the two cases of the rerooting process. For the first one, we have k possibilities. For the second, remark 5 implies that the black son of the root containing the old root is described by $zT^\circ(z)T^{k-1}(z)$ and the other black sons give the factor $\exp(zT^k(z))$:
$K^\circ(z) = kz^kT(z)+z^k\exp(zT^k(z))\, zT^\circ(z)T^{k-1}(z) = kz^kT(z)+z^{k+1}T^k(z)T^\circ(z)$.

Theorem 4. *There is a bijection between rerooted k-trees and pointed k-trees.*

Proof. We show that the generating functions for both classes are the same. We first need to express $T^\circ(z)$: adding a variable x to count white nodes for trees in $\mathcal{T}$, deriving with respect to x and setting $x = 1$, we get

$$T^\circ(z) = T(z) + k\exp(zT^k(z))zT^\circ(z)T^{k-1}(z) = T(z) + zkT^k(z)T^\circ(z),$$

which can also be obtained with a combinatorial argument: marking a white node consists either in marking the root or in choosing one of the sets of black nodes below it and marking in one of the k subtrees below this black node.

We now need to show the equality of generating functions $T^\bullet(z) = zT^k(z)$ $T^\circ(z)$, where $T^\bullet(z)$ counts trees in $\mathcal{T}$ with a mark on a black node.

For that, we use the fact that a tree in $\mathcal{T}$ with n black nodes contains $kn+1$ white nodes and that $[z^n]T^\bullet(z) = n[z^n]T(z)$. We thus have

$$\begin{aligned}[z^n]zT^k(z)T^\circ(z) &= \frac{1}{k}[z^n](T^\circ(z) - T(z)) = \frac{kn+1}{k}[z^n]T(z) - \frac{1}{k}[z^n]T(z)\\ &= n[z^n]T(z) = [z^n]T^\bullet(z).\end{aligned}$$

Hence $K^\bullet(z) = kz^kT(z) + z^kT^\bullet(z) = kz^kT(z) + z^{k+1}T^k(z)T^\circ(z) = K^\circ(z)$.

4.2 Limiting Distribution

We now come to study the distances to the first vertex of the root in a tree of type $\mathcal{K}^\circ$, which are exactly the distances to the marked vertex of a tree of type $\mathcal{K}^\bullet$. This allows to show that the distance between two random points in a random k-tree follows a Rayleigh distribution.

With similar notations as in section 3 and a similar analysis of the recursive structure of the trees, but more involved computations, we obtain:

Lemma 11. *The b.g.f. of vertices at distance d in K° can be expressed as*

$$K_1^\circ(z,u) = kz^k u^{k-1} T_{1,1}(z,u) + z^{k+1} u^k T_{1,1}^k(z,u) T^\circ(z),$$
$$K_d^\circ(z,u) = kz^k T_{d,1}(z,u) + z^{k+1} T_{d,1}^k(z,u) T_{d-1,k}^\circ(z,u).$$

Differentiating with respect to u gives

$$K_d^{\circ\prime}(z) = kz^k T^{k-1}(z) T_{d,1}'(z) + kz^{k+1} T^{k-1}(z) T_{d,1}'(z) T^\circ(z) + z^{k+1} T^k(z) T_{d-1,k}^{\circ\prime}(z).$$

In this function the dominant term is $z^{k+1}T^k(z)T_{d-1,k}^{\circ\prime}(z)$ and has the same singular behavior as $K_d'(z)$, up to a factor n, corresponding to the choice of a random vertex.

This result agrees with the general fact that in a very large random tree the root tends to have the same properties as any random vertex. We thus get exactly the same asymptotic distribution, as in section 3.

Theorem 5. *Given a random k-tree G over n vertices, the distance between two random vertices v, w of G has mean value of order $\sqrt{n}$ and is asymptotically Rayleigh distributed in the range $x\sqrt{n}$:*

$$\lim_{n\to\infty} \sqrt{n}\mathbb{P}(D(v,w) = \lfloor x\sqrt{n} \rfloor) = c_k^2 x e^{-\frac{(c_k x)^2}{2}}, \text{ with } c_k = k \sum_{i=1}^k \frac{1}{i}.$$

References

1. Arnborg, S., Proskurowski, A.: Linear time algorithms for NP-hard problems restricted to partial k-trees. Discrete Applied Mathematics 23(1), 11–24 (1989)
2. Beineke, L.W., Pippert, R.E.: The number of labeled k-dimensional trees. Journal of Combinatorial Theory 6(2), 200–205 (1969)
3. Bodini, O., Darrasse, A., Soria, M.: Distances in random apollonian network structures. In: FPSAC 2008. DMTCS Proceedings, pp. 307–318 (2008)
4. Drmota, M.: Random Trees: An Interplay between Combinatorics and Probability. Springer, Heidelberg (2009)
5. Flajolet, P., Sedgewick, R.: Analytic Combinatorics. Cambridge University Press, Cambridge (2009)
6. Fowler, T., Gessel, I., Labelle, G., Leroux, P.: The specification of 2-trees. Advances in Applied Mathematics 28(2), 145–168 (2002)
7. Ibarra, L.: The clique-separator graph for chordal graphs and subclasses of chordal graphs. In: Symposium on Discrete Mathematics, Nashville, TN (2004)

8. Klawe, M.M., Corneil, D.G., Proskurowski, A.: Isomorphism testing in hookup classes. SIAM Journal on Algebraic and Discrete Methods 3(2), 260–274 (1982)
9. Labelle, G., Lamathe, C., Leroux, P.: Labelled and unlabelled enumeration of k-gonal 2-trees. Journal of Combinatorial Theory, Series A 106(2), 193–219 (2004)
10. Markenzon, L., Justel, C.M., Paciornik, N.: Subclasses of k-trees: Characterization and recognition. Discrete Applied Mathematics 154(5), 818–825 (2006)
11. Meir, A., Moon, J.W.: On the altitude of nodes in random trees. Canadian Journal of Mathematics 30, 997–1015 (1978)
12. Moon, J.W.: The number of labeled k-trees. Journal of Combinatorial Theory 6(2), 196–199 (1969)
13. Proskurowski, A.: K-trees: representation and distances. In: Congressus Numerantium, vol. 29, pp. 785–794. Utilitas Mathematica (1980)
14. Proskurowski, A.: Recursive graphs, recursive labelings and shortest paths. SIAM Journal on Computing 10(2), 391–397 (1981)
15. Rose, D.J.: On simple characterizations of k-trees. Discrete Mathematics 7(3-4), 317–322 (1974)