SCALING LIMIT OF RANDOM $k$-TREES

MICHAEL DRMOTA AND EMMA YU JIN

Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Vienna, Austria

michael.drmota@tuwien.ac.at
yu.jin@tuwien.ac.at

Abstract. We consider a random $k$-tree $G_{n,k}$ that is uniformly selected from the class of labelled $k$-trees with $n + k$ vertices. Since 1-trees are just trees, it is well-known that $G_{n,1}$ (after scaling the distances by $1/(2\sqrt{n})$) converges to the Continuum Random Tree $\mathcal{T}_e$. Our main result is that for $k \neq 1$, the random $k$-tree $G_{n,k}$, scaled by $(k + 1)/(2\sqrt{n})$, converges to the Continuum Random Tree $\mathcal{T}_e$ too. In particular this shows that the diameter as well as the typical distance of two vertices in a random $k$-tree $G_{n,k}$ are of order $\sqrt{n}$.

1. Introduction and Main result

A $k$-tree is a generalization of a tree and can be defined recursively: a $k$-tree is either a complete $k$-graph on $k$ vertices (a $k$-clique) or a graph obtained from a smaller $k$-tree by adjoining a new vertex together with $k$ edges connecting it to a $k$-clique of the smaller $k$-tree (and thus forming a $(k + 1)$-clique), see Figure 2.1. In particular, a 1-tree is a usual tree. (Note that the parameter $k$ is always fixed.)

A $k$-tree is an interesting graph from an algorithmic point of view since many NP-hard problems on graphs have polynomial, in fact usually linear, dynamic programming algorithms when restricted to $k$-trees and their subgraphs for fixed values of $k$ [6, 40, 27]; subgraphs of $k$-trees are called partial $k$-trees. Such NP-hard problems include maximum independent set size, minimal dominating set size, chromatic number, Hamiltonian circuit, network reliability and minimum vertex removal forbidden subgraph [5, 9]. Several graphs which are important in practice [32], have been shown to be partial $k$-trees, among them are

1. Trees/ Forests (partial 1-trees)
2. Series parallel networks (partial 2-trees)
3. Outplanar graphs (partial 2-trees)
4. Halin graphs (partial 3-trees)

However, other interesting graph classes like planar graphs or bipartite graphs are not partial $k$-trees. $k$-trees are also very interesting from a combinatorial point of view. For example, the enumeration problem for $k$-trees has been studied in various ways, see [7, 34, 23, 13, 29, 30, 24, 25, 26]. The number of labelled $k$-trees has been determined by Beineke and Pippert [7], Moon [34], Foata [23], Darrasse and Soria [13]: as usual a $k$-tree on $n$ vertices is called labelled if the integers from $\{1, 2, \ldots, n\}$ have been assigned to its vertices (one-to-one) and two labelled $k$-trees are considered to be different if the corresponding edge sets are different. In order to analyze $k$-trees, it turns out that it is convenient to consider the number of hedra instead of the number of vertices as the size of a $k$-tree; we adopt the notions from [26]. A hedron is a $(k + 1)$-clique in a $k$-tree, and by definition a $k$-tree with $n$ hedra has $n + k$ vertices. A front of a $k$-tree is a $k$-clique. In what follows, we assume the $k$-trees are all labelled and a random $k$-tree with $n$ hedra is uniformly selected from the class of labelled $k$-trees with $n$ hedra.

Darrasse and Soria [13] showed a Rayleigh limiting distribution for the expected distance between pairs of vertices in a random $k$-tree, as it is known for usual trees and, thus, for 1-trees.

---

1 Corresponding author email: yu.jin@tuwien.ac.at; The first author is partially supported by the Austrian Science Fund FWF, Project F50-02. The second author is supported by the German Research Foundation DFG, JI 207/1-1, and the Austrian Research Fund FWF, project SFB F50 Algorithmic and Enumerative Combinatorics.
Inspired by this results, we expect that a random \( k \)-tree with \( n \) hedra, after scaling the distances to the root by \( 1/\sqrt{n} \), converges to the Continuum Random Tree multiplied by a scaling factor. For \( k = 1 \) this is true by a result of Aldous. Actually Aldous has proved in a series of seminal papers \([2, 3, 4]\) that a critical Galton-Watson tree conditioned on its size has the Continuum Random Tree (CRT) as its limiting object – and random 1-trees are a special case (with a Poisson offspring distribution). The concept Continuum Random Tree was also introduced by Aldous \([2, 3, 4]\) and further developed by Duquesne and Le Gall \([18, 19, 20]\).

Since Aldous’s pioneering work on the Galton-Watson trees, the CRT has been established as the limiting object of a large variety of combinatorial structures \([28, 38, 35, 36, 11, 31, 8, 12]\). A key idea in the study of these combinatorial objects is to relate them to trees endowed with additional structures by using an appropriate bijection. In the present case of \( k \)-trees, we encode them as so-called \( k \)-front coding trees via a bijection due to Darrasse and Soria in \([13]\), which was originally used to enumerate \( k \)-trees and to recursively count the distance between any two vertices in a random \( k \)-tree. Furthermore, in order to build a connection between the distance of two vertices in a random \( k \)-tree and the distance of two vertices in a critical Galton-Watson tree, we need to introduce the concept of a size-biased enriched tree. This is adapted from the size-biased Galton-Watson tree which was introduced by Addario-Berry, Devroye and Janson in \([1]\) and was further generalized to the size-biased \( \mathcal{R} \)-enriched trees by Panagiotou, Stufler and Weller in \([35, 36]\). Our enriched tree is slightly different to the size-biased \( \mathcal{R} \)-enriched tree and we use their ideas in \([38, 35]\) where an important step is to relate the distance between two vertices in a random graph to the distance between two blocks in a random size-biased \( \mathcal{R} \)-enriched tree.

Our main result establishes the convergence of a random \( k \)-tree to the CRT with respect to the Gromov-Hausdorff metric.

**Theorem 1.** Let \( G_{n,k} \) be the class of vertex labelled \( k \)-trees with \( n \) hedra and denote by \( G_{n,k} \) a random \( k \)-tree that is uniformly selected from the class \( G_{n,k} \) and by \( G_{n,k} \) a random \( k \)-tree that is rooted at a front. Then

\[
\frac{m_k}{2\sqrt{n}} G_{n,k} \xrightarrow{(d)} \mathcal{T}_\epsilon \quad \text{and} \quad \frac{m_k}{2\sqrt{n}} G_{n,k} \xrightarrow{(d)} \mathcal{T}_\epsilon
\]

with respect to the Gromov-Hausdorff metric, where \( m_k = k + 1 \) for \( k \neq 1 \) and \( m_1 = 1 \).

In particular this shows that the diameter as well as the typical distance of two vertices are of order \( \sqrt{n} \) and they have the same limiting distribution as random trees.

The plan of the paper is as follows. In Section 2 we recall the combinatorial background for \( k \)-trees, introduce the Boltzmann sampler – a method of generating efficiently a uniform random combinatorial object, describe Darrasse and Soria’s algorithm on computing the distances between two vertices in a \( k \)-tree, and present Aldous’s result on the convergence of critical Galton-Watson trees to the CRT \( \mathcal{T}_\epsilon \). In Section 3 we prove our main result – Theorem 1.

2. **Combinatorics, Boltzmann Sampler and Continuum Random Tree**

It was shown in \([7, 34, 23, 13]\) that the number \( L_k(n) \) of \( k \)-trees having \( n \) hedra is given by

\[
L_k(n) = \binom{n+k}{k}(kn+1)^{n-2}
\]

and, thus, asymptotically by \( L_k(n) \approx n^k(kn)^{n-2}/k! \) as \( n \to \infty \). Here we only review the generating function approach in \([13]\) to count \( L_k(n) \). The key ingredient to count the number \( L_k(n) \) in \([13]\) is a bijection between rooted \( k \)-trees and \( k \)-front coding trees rooted at a white node.

2.1. **A bijection.** A rooted \( k \)-tree is a \( k \)-tree rooted at a front (or equivalently a \( k \)-clique). A \( k \)-front coding tree is a bipartite tree of black and white nodes which is rooted at a white node and where every black node has precisely \( k \) successors. The bijection will be built in a way that black nodes in \( k \)-front coding trees correspond to hedra in \( k \)-trees. Every black node also gets a label which is equal to the label of one of the vertices of the corresponding hedron. A white node in a \( k \)-front coding tree corresponds to a front of the \( k \)-trees and is labelled by the set \( \{a_1, a_2, \ldots, a_k\} \) of labels of the corresponding front. A black node connects with a white node if the corresponding
bijection we mentioned above establishes a relation $G$ the class of one (black) child. Let $n$ coding trees with $k$ reduced to count the corresponding reduced $k$ (2.5) coding trees satisfy the following specification: $$n \text{-rooted } k \text{-front coding trees and } B_k \text{ be the class of } \circ \text{-rooted k-front coding trees. The bijection we mentioned above establishes a relation } G_k^\circ \simeq C_k. \text{ Furthermore, every } \circ \text{-rooted k-front coding tree can be identified as a set of } \circ \text{-rooted k-front coding trees and every } \circ \text{-front coding tree can be decomposed into a k-tuple of } \circ \text{-rooted k-front coding trees. Consequently, k-front coding trees satisfy the following specification:}$$

$$C_k = \text{SET}(B_k) \text{ and } B_k = \{\circ\} \ast \text{SEQ}_k(C_k).$$

In terms of exponential generating functions (where the size is always the number of black nodes), we thus get

$$C_k(x) = \exp(B_k(x)) \text{ and } B_k(x) = x \cdot C_k(x)^k.$$  

In particular $B_k(x)$ satisfies

$$B_k(x) = x \exp(kB_k(x)).$$

By applying the Lagrange inversion formula on (2.4), we obtain that the number of $\circ \text{-front coding trees with } n \text{ black nodes where the root } \circ \text{ has a fixed label } \{a_1, a_2, \ldots, a_k\}, \text{ is}$$

$$b_k(n) = n! [x^n] B_k(x) = (n - 1)! [x^{n-1}] \exp(knx) = (kn)^{n-1}$$

and the number of $\circ \text{-rooted k-front coding trees with } n \text{ black nodes where the root } \circ \text{ has a fixed label } \{a_1, a_2, \ldots, a_k\} \text{ is}$$

$$c_k(n) = n! [x^n] C_k(x) = (n - 1)! [x^{n-1}] \exp((kn + 1)x) = (kn + 1)^{n-1}.$$
Recall that the size of a $k$-front coding tree $T$, denoted by $|T|$, is the number of black nodes. Let $C_{n,k}$ be the class of $k$-front coding trees of size $n$ such that the white root has label $\{1, 2, \ldots, k\}$, then from (2.7) we find, the probability to uniformly choose a random rooted $k$-tree $C_{n,k}$ is equal to the probability to uniformly choose a $k$-front coding tree $C_{n,k}$ from the class $C_{n,k}$.

Since $C_{k}$ has a proper recursive specification (2.3), these random objects can be constructed (or sampled) by a so-called Boltzmann sampler $\Gamma C_{k}(x)$.

2.2. Boltzmann Sampler. Boltzmann samplers provide a way to efficiently generate a combinatorial object at random. They were introduced by Duchon, Flajolet, Louchard and Schaeffer [17] and were further developed by Flajolet, Fusy and Pivoteau [22]. Here we refer the readers to their papers [17, 22] for a detailed description of the Boltzmann samplers. An important property of Boltzmann samplers is that they generate objects of a given size uniformly.

More precisely we will describe a Boltzmann sampler $\Gamma C_{k}(x)$ with parameter $x = \rho_{k} = (ek)^{-1}$ (which is possible since $\lambda_{k} = B_{k}(\rho_{k}) = k^{-1} < \infty$).

**Lemma 2.** The following recursive procedure $\Gamma C_{k}(\rho_{k})$ terminates almost surely and draws a random $o$-rooted $k$-front coding tree according to the Boltzmann distribution with parameter $\rho_{k}$, i.e., any $o$-rooted $k$-front coding tree of size $n$ in the class $C_{n,k}$ is drawn with probability $\rho_{k}^{n}/(n! C_{k}(\rho_{k}))$.

\[ \Gamma C_{k}(\rho_{k}): \begin{align*}
    &x_{1} \leftarrow \text{a white node } o \\
    &m \leftarrow \text{Pois}(\lambda_{k}) \\
    &\text{for } i := 1 \rightarrow m \\
    &\quad \quad x_{2} \leftarrow \text{a single black node } * \\
    &\quad \quad \text{merge } x_{2} \text{ into } x_{1} \text{ by adding an edge } * - o \\
    &\quad \quad \mathcal{F} \leftarrow \text{an } m\text{-tuple } (\Gamma C_{k}(\rho_{k}), \ldots, \Gamma C_{k}(\rho_{k})) \\
    &\quad \quad \text{merge } \mathcal{F} \text{ into } x_{1} \text{ by connecting } x_{2} \text{ to the roots of } \mathcal{F} \\
    &\quad \quad x_{1} \leftarrow \text{label the black nodes of } x_{1} \text{ uniformly at random} \\
    &\text{return } x_{1}
\end{align*} \]

Note that $k$-front coding trees satisfy the specification (2.3), but they do not represent the distance relation in the $k$-trees. See Figure 2.1. Since we have fixed the label on the white root $o$, which is $\{1, 2, \ldots, k\}$, the labels on the black nodes of $\Gamma C_{k}(\rho_{k})$ determine the corresponding labels on the white nodes.

2.3. $k$-tree distance algorithm. For a random $k$-front coding tree $C_{n,k}$ of size $n$, let $M_{n,k}$ be the corresponding $k$-tree under the bijection $C_{n,k} \rightarrow M_{n,k}$ in subsection 2.1, then $M_{n,k}$ is rooted at a front $\{1, 2, \ldots, k\}$. We use the notation $(i^{m}, j^{k-m})$ to represent the sequence $(i, i, \ldots, i, \ldots, j, \ldots, j)$ of length $k$ that has $m$ occurrences of $i$ and $(k - m)$ occurrences of $j$. Here we shall consider the distances to the vertex 1 in a $k$-tree $M_{n,k}$. Darrasse and Soria [13] provided an algorithm to calculate the distances to the vertex 1 in a $k$-tree $M_{n,k}$ by marking the distances on the corresponding $k$-front coding tree $C_{n,k}$, which is similar to the algorithm given by Proskurowski in [37]. Note that every black node of the $k$-front coding tree is related to a vertex of the corresponding $k$-tree via its label, and the vertices that label a white node of the $k$-front tree represent $k$ vertices that constitute a front of the corresponding $k$-tree. We will recall Darrasse and Soria’s algorithm.

**Algorithm 1:** Distances in a $k$-tree

Input: a $k$-front coding tree $C$ and a sequence $(a_{i})_{i=1}^{k} = (0, 1^{k-1})$

Output: an association table (vertex, distance)

\[ p := \min\{a_{i}\}_{i=1}^{k} + 1 \text{ and } A = \emptyset \]

for all sons $v$ of the root $C$ do

\[ A := A \cup \{(v, p)\} \]

for $i := 1 \rightarrow k$ do

\[ A := A \cup \text{the recursive call on the } i\text{-th son of } v \text{ and } (a_{1}, \ldots, a_{i-1}, p, a_{i+1}, \ldots, a_{k}) \]

return $A$

If we implement this algorithm on the 2-front coding tree (middle) in Figure 2.1, we get a distance table marked on every black node in Figure 2.2. The distance sequences on the white nodes help us to recursively mark the distances on the black nodes.
2.4. Gromov-Hausdorff convergence and the CRT. Let $e = (e_s)_{0 \leq s \leq 1}$ denote the Brownian excursion of duration one. Then this (random) continuous function $e$ induces a pseudo-metric on the interval $[0, 1]$ by

$$d_e(u, v) = e(u) + e(v) - 2 \inf_{u \leq s \leq v} e(s)$$

for $u \leq v$. This defines a metric on the quotient $\mathcal{T}_e = [0, 1]/\sim$ where $u \sim v$ if and only if $d_e(u, v) = 0$. The corresponding random pointed metric space $(\mathcal{T}_e, d_e, r_0(\mathcal{T}_e))$, where $r_0(\mathcal{T}_e)$ is the equivalence class of the origin, is the Continuum Random Tree (CRT). We will simply use $\mathcal{T}_e$ to denote the CRT. (Recall that the isometry classes of (pointed) compact metric spaces $\mathcal{K}^*$, where a pointed space is a triple $(X, d, r)$, where $(X, d)$ is a metric space and $r \in X$ is a distinguished element, constitute a Polish space with respect to the (pointed) Gromov-Hausdorff metric $d_{\text{GH}}$. We refer the readers to [10, 21] for a full description of this metric.)

Let $T_n$ be a Galton-Watson tree conditioned on having $n$ vertices, $T_n$ is critical if the offspring distribution $\xi$ of $T_n$ satisfies $\mathbb{E}(\xi) = 1$. $T_n$ is aperiodic if $\gcd\{j : \mathbb{P}(\xi = j) > 0\} = 1$, see [15]. The convergence of $T_n$ (properly scaled) to $\mathcal{T}_e$ is due to Aldous [4].

**Theorem 3.** Let $T_n$ be a Galton-Watson tree conditioned on having $n$ vertices, where the offspring distribution $\xi$ of $T_n$ is aperiodic, critical and has finite variance $\text{Var}(\xi) = \sigma^2$. As $n$ tends to infinity, $T_n$ with edges rescaled to length $\sigma/(2\sqrt{n})$ converges in distribution to the CRT, i.e.,

$$\frac{\sigma}{2\sqrt{n}}T_n \xrightarrow{(d)} \mathcal{T}_e \quad \text{in the metric space} \ (\mathcal{K}^*, d_{\text{GH}}).$$

The Galton-Watson tree conditioned on having $n$ vertices is also called the conditioned Galton-Watson tree. The conditioned Galton-Watson trees are essentially the same as the random simply generated trees, see [14, 15].

3. Proof of the Main result

Let $C_{n,k}$ denote the the random $k$-front coding tree that is generated by the Boltzmann sampler $\Gamma C_k(\rho_k)$ having exactly $n$ black nodes. Furthermore let $M_{n,k}$ be the corresponding random $k$-tree under the bijection $C_{n,k} \mapsto M_{n,k}$. Finally let $T_{n,k}$ be the reduced tree obtained from $C_{n,k}$ by replacing every edge $\bullet \rightarrow \circ \rightarrow \bullet$ that passes a non-root white node, by an edge $\bullet \rightarrow \bullet$ and removing all the white-node leaves, see Figure 2.1 for an example.

From the construction of the Boltzmann sampler $\Gamma C_k(\rho_k)$, it is clear that $T_{n,k}$ is a Galton-Watson tree conditioned on having exactly $n$ black nodes where the offspring $\xi$, of the white root node is Poisson distributed with parameter $k^{-1}$ whereas the offspring $\xi$ of every black node is Poisson distributed with parameter $k\lambda_k = kB_k(\rho_k) = 1$, that is, the probability that a black node
has m offsprings in \( T_{n,k}^0 \) is
\[
\mathbb{P}(\xi = m) = \exp(-k\lambda_k) \cdot \frac{(k\lambda_k)^m}{m!} = \frac{\exp(-1)}{m!} \quad \text{and} \quad \mathbb{E}\xi = 1.
\]

We can modify the tree \( T_{n,k}^0 \) into a conditioned critical Galton-Watson tree \( T_{n,k} \) by replacing the offsprings of the white root also by a Poisson distribution with parameter 1.

For any two black nodes \( x, y \) in \( C_{n,k} \), we set \( d_{C_{n,k}}(x, y) = \text{dist}_{T_{n,k}}(x, y) \), where dist denotes the usual graph theoretical distance. For any block of type \( a \), either a black node or a white node in the same block of \( x, y \) is different to the distance \( \text{dist}_{M_{n,k}}(x, y) \) in the original \( k \)-tree \( M_{n,k} \). In order to represent the distances \( \text{dist}_{M_{n,k}}(x, y) \) for any two nodes \( x, y \) in the tree \( C_{n,k} \), we need to decompose \( C_{n,k} \) into blocks according to the distance table from Algorithm 1. We implement the Algorithm 1 on the random tree \( C_{n,k} \) to have every black node marked with a distance and every white node marked with a distance sequence. For this random tree \( C_{n,k} \), denote by \( C_{i,n,k} \) a subtree of \( C_{n,k} \) that we call block of type \( i \):

1. \( C_{i,n,k} \) is rooted at the root and is induced by the root and all the black nodes that are in distance one to the vertex 1.
2. \( C_{i,n,k} \), \( i \geq 2 \), is rooted at a white node with distance sequence \( ((i - 1)\delta) \) and is induced by this node and all its black descendants that have distance \( i \) to the vertex 1.

By construction, there is only one subtree \( C_{1,n,k} \) in \( C_{n,k} \), but there could be many subtrees \( C_{i,n,k} \) of \( C_{n,k} \) for \( i \neq 1 \), see Figure 3.1. For any two black nodes \( x, y \) in \( C_{n,k} \), let \( \delta_{C_{n,k}}(x, y) = a - 1 \) where \( a \) is the minimal number of blocks necessary to cover the path connecting \( x \) and \( y \). In particular if \( x, y \) are in the same block of \( C_{n,k} \), then \( \delta_{C_{n,k}}(x, y) = 0 \). The following lemma will show, for any two black nodes \( x, y \), the distance \( \text{dist}_{M_{n,k}}(x, y) \) is almost the same as the block-distance \( \delta_{C_{n,k}}(x, y) \).

![Figure 3.1. A decomposition of a random 2-front coding tree \( C_{n,2} \) into blocks \( C_{i,n,2} \) (left) where the pair \((a, b)\) of integers represents the distance sequence on the root of a block. A spine (right) consists of selected good nodes in \( C_{n,2} \).](image.png)

**Lemma 4.** Let \( C_{n,k} \) denote the tree corresponding to the Boltzmann sampler \( \Gamma_{\rho_k}(\rho_k) \) conditioned on having \( n \) black nodes, let \( M_{n,k} \) be the corresponding \( k \)-tree under the bijection \( C_{n,k} \mapsto M_{n,k} \). Then for any two black nodes \( x, y \) in \( M_{n,k} \),
\[
\text{dist}_{M_{n,k}}(x, y) = \delta_{C_{n,k}}(x, y) + i \quad \text{where} \quad i \in \{0, 1, 2, 3\}.
\]

**Proof.** If \( x, y \) are in the same block, i.e., \( \delta_{C_{n,k}}(x, y) = 0 \). If both of them are in a block \( C_{i,1,n,k} \), then \( \text{dist}_{M_{n,k}}(x, y) \leq \text{dist}_{M_{n,k}}(x, 1) + \text{dist}_{M_{n,k}}(y, 1) = 2 = \delta_{C_{n,k}}(x, y) + 2 \). If both of them are in a block \( C_{i+1,1,n,k} \) for some \( i \geq 1 \), recall that the root of \( C_{i+1,1,n,k} \) is a white node with distance sequence \((i\delta)\). Suppose the root of \( C_{i+1,1,n,k} \) has label \( \{a_1, a_2, \ldots, a_k\} \), then for \( x \in C_{i+1,1,n,k} \), there exists an integer \( p \) such that \( \text{dist}_{M_{n,k}}(a_p, x) = 1 \). Otherwise if for all \( m \leq k \), \( \text{dist}_{M_{n,k}}(a_m, x) > 1 \). It follows that \( \text{dist}_{M_{n,k}}(x, 1) > i + 1 \), which contradicts with the fact \( x \in C_{i+1,1,n,k} \). Similarly, there is an integer \( q \) such that \( \text{dist}_{M_{n,k}}(a_q, y) = 1 \). Consequently \( \text{dist}_{M_{n,k}}(x, y) \leq \text{dist}_{M_{n,k}}(a_p, x) + \text{dist}_{M_{n,k}}(a_q, y) + \text{dist}_{M_{n,k}}(a_p, a_q) = 3 \), which implies (3.2).

If \( x, y \) are not in the same block, let \( b \) be the lowest common parent of \( x \) and \( y \) in \( C_{n,1} \), then \( b \) is either a black node or \( b \) is the white root. If \( b \) is a black node, let \( a_1 \) (resp. \( b_1 \)) be the second black node on the path \( b - o - a_1 - \cdots - o - x \) (resp. \( b - o - b_1 - \cdots - o - y \)) in \( C_{n,k} \). If \( b \) is the root, let
Lemma 5.

Proof. Let $\xi_{k,i}$ be the random variable counting the number of good black nodes $v$ in a block $B_{i,k}$ of type $i$ in $C_k$. Then $E[\xi_{k,i}] = 1$.

Proof. The offspring $\eta$ of the white root in $C_k$ is Poisson distributed with parameter $k^{-1}$ and the offspring of every black node in $C_k$ is distributed as the sum of $k$ independent and identically distributed random variables $\eta_1, \eta_2, \ldots, \eta_k$ where each $\eta_i$ is a copy of $\eta$. The distance sequence on every white node of $C_k$ determines if his children (black nodes) are good or not. We first compute $E[\xi_{k,1}]$. The white root of $B_{1,k}$ has distance sequence $(0, 1^{k-1})$ and all his children are good black nodes. So the first generation has $E(\eta) = k^{-1}$ expected number of good black nodes. Assume $\mu_1$ is a good black node in the first generation, $\mu_1$ has $k$ white-node children in $C_k$, among which $(k-1)$ have distance sequence $(0, 1^{k-1})$ and they have $1 - k^{-1}$ expected number of good black-node children in $C_k$. By repeating this process on these good black-node children, we get

$$E(\xi_{k,1}) = \frac{1}{k} + \frac{k-1}{k} \left( \frac{1}{k} + \frac{k-1}{k} + \cdots \right) = \frac{1}{k} \sum_{i=0}^{k-1} \left( \frac{k-1}{k} \right)^i = 1.$$ 

For $i \neq 1$, we can compute $E(\xi_{k,i})$ by repeating the same procedure as that for $E(\xi_{k,1})$, namely the expected number $E(\xi_{k,i})$ of good black nodes in a block $B_{i,k}$ is equal to the expected number of good black nodes in its sub-rooted block at a black node $\omega_i$, multiplied by $k^{-1}$, where $\omega_i$ is the first good black node on its path to the root of block $B_{i,k}$, from which it follows the expected number of good descendents of $\omega_i$ is $k$. This implies $E(\xi_{k,i}) = k^{-1} \cdot k = 1$. □
Figure 3.2. A 2-front coding tree $C_2$ with good nodes drawn in big $\bullet$ (left) and a size-biased enriched tree $\hat{C}_2^{(3)}$ with a spine consisting of selected good nodes drawn in big $\blacksquare$ (right).

We will next define a size-biased enriched tree $\hat{C}_k^{(m)}$ from a random $k$-front coding tree $C_k$. This construction comes from [1], which is a truncated version of the infinite size-biased Galton-Watson tree introduced by Lyons, Pemantle and Peres [33]. Let $\xi_{k,i}$ be a random variable with the size-biased distribution

$$P(\xi_{k,i} = q) = q \cdot P(\xi_{k,i} = q).$$

The expected value $E(\xi_{k,i}) = 1$ in Lemma 5 guarantees that $\xi_{k,i}$ is a probability distribution on $\mathbb{N}^+ = \{1, 2, \ldots\}$.

The size-biased enriched tree $\hat{C}_k^{(m)}$ is now defined as follows. It starts with a mutant block $B_{1,k}$ which is rooted at a usual root (that has distance sequence $(0, 1^k)$) and contains good nodes. We now choose one of these good nodes (which number is distributed according to $\xi_{k,1}$) and call it heir (and also mutant). The block $B_{2,k}$ that is rooted at the child with distance sequence $(1^k)$ of this heir will be the next mutant block, where we again assume that it has at least one good node. All other blocks that are adjacent to $B_{1,k}$ are normal. We again choose one of the good nodes of the mutant block $B_{2,k}$ (which number is distributed according to $\xi_{k,2}$) and proceed inductively to choose mutant blocks and heirs till $B_{m,k}$. All other blocks stay normal. We denote the heir in the $m$-th mutant block $B_{m,k}$ by $h$. The path from the root to $h$ is called spine of $\hat{C}_k^{(m)}$, see Figure 3.2.

The probability that a given mutant block contains $q$ good nodes and one of them is chosen as heir is, see (3.3), $q^{-1}P(\hat{\xi}_{k,i} = q) = P(\xi_{k,i} = q)$. For any given random $k$-front coding tree $T$, let $T^\alpha$ denote the tree $T$ with a fixed spine $\alpha$ of block-depth $m$. Then the probability

$$P(\hat{C}_k^{(m)} = T^\alpha, \text{ with } \alpha \text{ as spine}) = P(C_k = T).$$

This shows, once the spine is fixed, that the probability that the size biased tree $\hat{C}_k^{(m)}$ equals $T^\alpha$ is the same as the probability of generating $T$. In fact, (3.4) is true for any spine $\alpha$, see Eq.(3.2) in [1]. We will need (3.4) to build a connection between $m_k \delta_{C_{n,k}}(x, y)$ and $d_{C_{n,k}}(x, y)$ with high probability in Lemma 6.

Lemma 6. Let $C_{n,k}$ be the class of rooted $k$-front coding trees of size $n$ such that the white root has label $\{1, 2, \ldots, k\}$ and $C_{n,k} \in C_{n,k}$ is uniformly selected at random. Let $m_k = k + 1$ for $k \geq 2$ and $m_1 = 1$. Then for all $s > 1$ and $0 < \epsilon < 1/2$ with $2s\epsilon > 1$, we have for all black nodes $x, y$ in
\( C_{n,k} \) such that \( x \) is an ancestor of \( y \), that one of these two properties
\[
(3.5) \quad \delta_{c_{n,k}}(x,y) \geq \log^2(n) \quad \text{and} \quad |d_{c_{n,k}}(x,y) - m_k \delta_{c_{n,k}}(x,y)| \leq \delta_{c_{n,k}}(x,y)^{1/2+\epsilon},
\]
\[
(3.6) \quad \delta_{c_{n,k}}(x,y) < \log^2(n) \quad \text{and} \quad d_{c_{n,k}}(x,y) \leq \log^{2+2\epsilon}(n)
\]
holds with high probability.

**Proof.** Suppose the opposite of (3.5) is true, that is, there exist black nodes \( x, y \) in \( C_k \) such that \( x \) is an ancestor of \( y \) and they satisfy
\[
(3.7) \quad \delta_{c_{i,k}}(x,y) \geq \log^2(|C_k|) \quad \text{and} \quad |d_{c_{i,k}}(x,y) - m_k \delta_{c_{i,k}}(x,y)| > \delta_{c_{i,k}}(x,y)^{1/2+\epsilon}.
\]
We will denote by \( F_1 \) the set of triples \((C_k, x, y)\) (with \( x, y \) in \( C_k \)) that satisfy (3.7). Thus we just have to show that \( \mathbb{P}((C_k, x, y) \in F_1 | |C_k| = n) = o(1) \) as \( n \) tends to infinity.

Recall that \( C_{n,k} \) is the set of \( k \)-front coding trees generated by the Boltzmann sampler \( \Gamma C_k(\rho_k) \) with \( n \) black nodes. Thus, from Lemma 2 it follows that
\[
\mathbb{P}(C_{n,k}) = \mathbb{P}(|\Gamma C_k(\rho_k)| = n) = \frac{c_k(n)\rho_k^n}{n! C_{k}(\rho_k)} \cdot \frac{(kn + 1)^{n-1}}{n! \exp(k-1)} \cdot \frac{1}{\sqrt{2\pi n}} \sim \frac{1}{k^{3/2}} n^{-3/2} \quad \text{as} \quad n \to \infty.
\]
We apply (3.4) on the random \( k \)-front coding tree \( C_{n,k} \) with a spine that connects \( x \) to \( y \). The block-depth of this spine is at least \( \log^2 n \) by assumption (3.7), which leads to
\[
\mathbb{P}((C_k, x, y) \in F_1 | |C_k| = n) \leq \mathbb{P}(C_{n,k})^{-1} \sum_{m=\log^2 n}^{n-1} \mathbb{P}((\hat{C}^{(m)}_k, x, y) \in F_1 \text{ and } |\hat{C}^{(m)}_k| = n)
\]
\[
= \frac{k^{3/2} \pi}{\exp(-k^{-1})} n^{3/2} \sum_{m=\log^2 n}^{n-1} \mathbb{P}((\hat{C}^{(m)}_k, x, y) \in F_1 \text{ and } |\hat{C}^{(m)}_k| = n).
\]

Here the length of the spine in \( \hat{C}^{(m)}_k \) is distributed as the sum of \( m \) independent random variables \( \zeta_{1,k}, \zeta_{2,k}, \ldots, \zeta_{m,k} \) where each \( \zeta_{i,k} \) is distributed as the length of the path from the selected good node in some block \( B_{i,k} \) to the root of this block. We have for \( k \geq 2 \),
\[
\mathbb{P}(\zeta_{i,k} = t) = \frac{1}{k} \cdot \left( 1 - \frac{1}{k} \right)^{t-2} \quad \text{where} \quad i \neq 1, t \geq 2
\]
\[
\mathbb{P}(\zeta_{1,k} = t) = \frac{1}{k} \cdot \left( 1 - \frac{1}{k} \right)^{t-1} \quad \text{where} \quad t \geq 1.
\]
(We just have to extend the proof idea of Lemma 5.) For the case \( k = 1 \), every \( \zeta_{i,1} \) is distributed with probability \( \mathbb{P}(\zeta_{i,1} = 1) = 1 \). As an immediate consequence, \( \zeta_{i,k} \) has finite exponential moments for every \( i, k \) and \( \mathbb{E}(\zeta_{i,k}) = k + 1 \), \( \mathbb{E}(\zeta_{1,k}) = k \) for \( k \geq 2, i \neq 1 \). For the case \( k = 1 \) we have \( \mathbb{E}(\zeta_{i,1}) = 1 \) for every \( i \). We set \( m_k = k + 1 \) for \( k \geq 2 \) and \( m_1 = 1 \). Furthermore, the assumption in (3.7) implies
\[
(3.9) \quad \mathbb{P}((\hat{C}^{(m)}_k, x, y) \in F_1 \text{ and } |\hat{C}^{(m)}_k| = n) \leq \mathbb{P}(\sum_{i=1}^{m} \zeta_{i,k} - m \cdot m_k > m^{1/2+\epsilon}).
\]
By applying the deviation inequality (see [1, 35, 36]) on the random variables \( \zeta_{1,k}, \zeta_{2,k}, \ldots, \zeta_{m,k}, \) we get for some positive constant \( c_1 \),
\[
\mathbb{P}(\sum_{i=1}^{m} \zeta_{i,k} - m \cdot m_k > m^{1/2+\epsilon}) \leq 2 \exp(-c_1 (\log n)^{2\epsilon}) = o(n^{-5/2}).
\]
Together with (3.8) and (3.9), we can conclude that \( \mathbb{P}((C_k, x, y) \in F_1 | |C_k| = n) = o(1) \).

Now we turn to suppose the opposite of (3.6) is true, i.e., there exist black nodes \( x, y \) in \( C_k \) such that \( x \) is an ancestor of \( y \). They satisfy
\[
(3.10) \quad \delta_{C_k}(x,y) < \log^2(|C_k|) \quad \text{and} \quad d_{C_k}(x,y) > \log^{2+2\epsilon}(|C_k|).
\]
We use the notation \( F_2 \) to represent the set of triples \((C_k, x, y)\) (with \( x, y \) in \( C_k \)) that satisfy (3.10). Again from (3.8) and from the deviation inequality, we obtain for some positive constant \( c_2 \),
\[ P((C_k, x, y) \in \mathcal{F}_2 | |C_k| = n) \leq \frac{k \sqrt{2\pi}}{\exp(-k^{-1})} n^{3/2} \sum_{m=1}^{\log^* n} P((\hat{C}^{(m)}_k, x, y) \in \mathcal{F}_2 \text{ and } |\hat{C}^{(m)}_k| = n) \]

\[ \leq \frac{k \sqrt{2\pi}}{\exp(-k^{-1})} n^{3/2} \sum_{m=1}^{\log^* n} \left( \sum_{i=1}^{m} \zeta_{i,k} > \log^{+2} n + \exp(-c_2 \log^{2+4}(n)) = o(1) \right) \]

and the proof is complete. \(\Box\)

Now we are ready to prove our main result.

**Proof of Theorem 1.** It follows from Lemma 6 that with high probability
\[ |d_{C_n,k}(x, y) - m_k \delta_{C_n,k}(x, y)| \leq \delta_{C_n,k}(x, y)^{1/2+\epsilon} + \log^{+2}(n) \]
holds for all black nodes \(x, y\) where \(x\) is an ancestor of \(y\) in the random \(k\)-front coding tree \(C_{n,k}\).
For any two black nodes \(\mu, \nu\) in \(C_{n,k}\), let \(\alpha\) be the lowest common ancestor of \(\mu\) and \(\nu\) (\(\alpha\) could be the white root of \(C_{n,k}\)), then
\[ |d_{C_n,k}(\mu, \nu) - m_k \delta_{C_n,k}(\mu, \nu)| \leq \delta_{C_n,k}(\mu, \alpha)^{1/2+\epsilon} + \delta_{C_n,k}(\nu, \alpha)^{1/2+\epsilon} + 2 \log^{+2}(n) \]
\[ \leq 2D(C_{n,k})^{1/2+\epsilon} + 2 \log^{+2}(n) \]
where \(D(C_{n,k})\) is the diameter of random tree \(C_{n,k}\). We divide both sides of this inequality by \(\sqrt{n}\) and obtain
\[ (3.11) \quad \frac{|d_{C_n,k}(\mu, \nu) - m_k \delta_{C_n,k}(\mu, \nu)|}{\sqrt{n}} \leq \frac{2D(C_{n,k})^{1/2+\epsilon}}{\sqrt{n}} + \frac{2 \log^{+2}(n)}{\sqrt{n}}. \]

Recall that \(T_{n,k}\) contains the white root and all the black nodes of \(C_{n,k}\). \(T_{n,k}\) is a critical conditioned Galton-Watson tree. The only difference between \(C_{n,k}\) and \(T_{n,k}\) is that the offspring of the white root in \(C_{n,k}\) is Poisson distributed with parameter \(k^{-1}\), while the offspring of the white root in \(T_{n,k}\) is Poisson distributed with parameter 1. If both black nodes \(\mu, \nu\) are contained in a random \(k\)-front coding tree \(C_{n,k}\), then they must be in the random tree \(T_{n,k}\), which indicates \(d_{C_n,k}(\mu, \nu) = \text{dist}_{T_{n,k}}(\mu, \nu)\) and \(D(C_{n,k}) \leq D(T_{n,k})\). Consequently, (3.11) rewrites to
\[ \frac{|\text{dist}_{T_{n,k}}(\mu, \nu) - m_k \delta_{C_n,k}(\mu, \nu)|}{\sqrt{n}} \leq \frac{2D(T_{n,k})^{1/2+\epsilon}}{\sqrt{n}} + \frac{2 \log^{+2}(n)}{\sqrt{n}}. \]

The diameter \(D(T_{n,k})\) of a random tree \(T_{n,k}\) is less than the height \(H(T_{n,k})\) of \(T_{n,k}\) multiplied by \(2\). By applying the tails for the height of \(T_{n,k}\), see Theorem 1.2 in [1] and left-tail upper bounds for the height in [1], we obtain for the Gromov-Hausdorff distance
\[ d_{GH}\left( T_{n,k}, \frac{m_k C_{n,k}}{\sqrt{n}} \right) \leq \max_{\mu, \nu} \left| \frac{\text{dist}_{T_{n,k}}(\mu, \nu)}{\sqrt{n}} - \frac{m_k \delta_{C_n,k}(\mu, \nu)}{\sqrt{n}} \right| \overset{p}{\rightarrow} 0. \]

Namely, for any fixed \(\varepsilon\), the probability of the event \(d_{GH}\left( T_{n,k}, \frac{m_k C_{n,k}}{\sqrt{n}} \right) \leq \varepsilon\) converges to 1 as \(n\) tend to infinity.

Since the variance of \(\varepsilon\), the probability of the offspring distribution in the random tree \(T_{n,k}\) is 1, it follows from Theorem 3 that \(T_{n,k}/(2\sqrt{n}) \overset{(d)}{\rightarrow} T_\varepsilon\). Hence we get \(m_k C_{n,k}/(2\sqrt{n}) \overset{(d)}{\rightarrow} T_\varepsilon\) and consequently with the help of Lemma 4, we have
\[ \frac{m_k M_{n,k}}{2\sqrt{n}} \overset{(d)}{\rightarrow} T_\varepsilon. \]
Let \(G_{n,k}\) be the class of \(k\)-trees with \(n\) hexdra, let \(G_{n,k}\) be the random \(k\)-tree that is uniformly selected from the class \(G_{n,k}\). Let \(G_{n,k}^0\) be the random \(k\)-tree \(G_{n,k}\) rooted at a front. Then the probability to uniformly choose \(G_{n,k}\) from the class \(G_{n,k}\) is equal to the probability to uniformly choose \(C_{n,k}\)
from the class $\mathcal{C}_{n,k}$, namely the rooting and labeling process will not affect the probability space. This indicates,

$$\frac{m_k G_{n,k}}{2\sqrt{n}} \xrightarrow{d} \mathcal{T}_e \quad \text{and} \quad \frac{m_k G^*_{n,k}}{2\sqrt{n}} \xrightarrow{d} \mathcal{T}_e$$

where $m_k = k + 1$ for $k \neq 1$ and $m_1 = 1$. \hfill \Box

REFERENCES


[38] B. Stufler, The CRT is the scaling limit of unlabelled unrooted trees, submitted.