A DISTRIBUTIONAL STUDY OF THE PATH EDGE-COVERING NUMBERS FOR RANDOM TREES

ALOIS PANHOLZER

Abstract. We study for various tree families the distribution of the number of edge-disjoint paths required to cover the edges of a random tree of size $n$. For all tree families considered we can show a central limit theorem with expectation $\sim \mu n$ and variance $\sim \nu n$ with constants $\mu, \nu$ depending on the specific tree family.

1. Introduction

A set $P$ of edge-disjoint paths in a graph $G$ is called an edge-covering of $G$ if every edge of $G$ is contained in a path of $P$. The path edge-covering number of $G$ is then defined as the smallest number $p(G)$ of paths in such an edge-covering of $G$. This parameter appeared in the study of certain models of information retrieval structures (see, e.g., [6]).

If we are considering a tree $T$, then it is well-known (see [12]) that the path edge-covering number of $T$ is given by

$$p(T) = \frac{1}{2}X(T),$$

where $X(T)$ denotes the number of nodes in $T$ with odd degree. See Figure 1 for an example of a minimal edge-covering of a tree. Equation (1) can be shown easily via induction on the size $|T|$ of the tree $T$. We first remark that the inequality $P(T) \geq \frac{1}{2}X(T)$ holds due to the fact that when considering an arbitrary edge-covering $P(T)$ we get that for every node $x \in T$ with odd degree there must be at least one path in $P(T)$ with $x$ as an endpoint. Thus it suffices to show that there always exists an edge-covering with $\frac{1}{2}X(T)$ nodes. If $|T| = 1$ then $T$ is a single node and equation (1) holds. Now we assume that (1) holds for all trees $T'$ with $|T'| < n$, for a given $n \geq 2$. To show this equation also for trees $T$ of size $n$ we take two endnodes $y$ and $z$ of $T$ and consider the path $Q = yx_1 \ldots x_{r-1}z$, $r \geq 1$, which connects $y$ and $z$. After removing all edges from $T$ contained in $Q$ we obtain two isolated nodes $y$ and $z$ and a forest of trees $T_1, \ldots, T_{r-1}$. Since $|T_i| < n$, for $1 \leq i \leq r - 1$, we obtain due to the induction hypothesis $p(T_i) = \frac{1}{2}X(T_i), 1 \leq i \leq r - 1$. Thus we have shown existence of an edge-covering of $T$ with $1 + \sum_{i=1}^{r-1} p(T_i) = 1 + \frac{1}{2} \sum_{i=1}^{r-1} X(T_i) = \frac{1}{2}X(T)$ nodes, which proves equation (1) also for trees $T$ of size $n$.

Date: December 4, 2006.

This author’s work was supported by the Austrian Science Foundation FWF, grant S9608-N13.

An extended abstract of this paper was presented at GRACO 2005 (2nd Brazilian Symposium on Graphs, Algorithms, and Combinatorics) and appeared, under a different title, in Electronic Notes in Discrete Mathematics 19 (2005), 163–169.
Figure 1. A tree $T$ of size 17 with 14 nodes of odd degree; it has thus a path edge-covering number $p(T)$ of 7. An example of a minimal edge-covering is given.

Meir and Moon [8] studied the average behaviour of the parameter $p(T)$ for random trees of size $n$ of various tree families of interest: if we denote by $P_n$ the random variable that counts the path edge-covering number $p(T)$ of a randomly chosen tree $T$ of size $n$ for the tree family considered, then it was computed therein the path edge-covering constant $\mu := \lim_{n \to \infty} \frac{E(P_n)}{n}$ for simply generated trees, recursive trees, rooted and unrooted unlabelled trees. As the underlying model of randomness it was always used the random tree model: for unweighted trees it is assumed that every tree of size $n$ in the tree family considered can be selected with equal probability, whereas for weighted trees it is assumed that every tree of size $n$ in the tree family considered can be selected with probability proportional to its weight.

The intention of the present work is to extend the studies of [8] of the average path edge-covering number for random trees, where we are now interested in the distributional behaviour of $P_n$. It turns out that, for the tree families considered there (and also for two further tree families studied in the present paper), $P_n$ is asymptotically Gaussian distributed with expectation $E(P_n) \sim \mu n$ and variance $\nu(P_n) \sim \nu n$, with constants $\mu, \nu$ depending on the specific tree family.

2. Preliminaries

2.1. Tree families considered.

2.1.1. Simply generated tree families. They were introduced in [9] and include several important tree families, e. g., binary trees, unordered labelled trees (= Cayley trees) and planted plane trees (= ordered trees). Moreover, they are strongly related to Galton-Watson branching processes, since it is well known that random simply generated trees are essentially the same as conditioned Galton-Watson trees, obtained as the family tree of a Galton-Watson process conditioned on the given total size.
A sequence of non-negative real numbers \((\varphi_k)_{k \geq 0}\) with \(\varphi_0 > 0\) (the degree-weight sequence) is used to define the weight \(w(T)\) of any planted plane tree \(T\) by \(w(T) := \prod_v \varphi_{d^+(v)}\), where \(v\) ranges over all vertices of \(T\) and \(d^+(v)\) is the out-degree of \(v\). The resulting simply generated tree family consists then of all trees \(T\) with \(w(T) \neq 0\) together with its weights \(w(T)\). It follows further that the generating function \(T(z) = \sum_{n \geq 0} T_n z^n\) of the quantity \(T_n := \sum_{|T|=n} w(T)\) (with \(|T|\) the size of the tree \(T\)) satisfies the functional equation
\[
T(z) = z \varphi(T(z)),
\]
where the degree-weight generating function \(\varphi(t)\) is given by \(\varphi(t) = \sum_{k \geq 0} \varphi_k t^k\).

As sufficient conditions on the validity of our results given in Section 3, we will only make few restrictions on the degree-weight generating functions \(\varphi(t)\): (i) we will suppose, that \(\varphi(t)\) has a positive radius of convergence \(R > 0\), (ii) we will assume, that there exists a minimal positive solution \(\tau < R\) of the equation \(t \varphi'(t) = \varphi(t)\), (iii) we will further assume, that the period \(q\) of \(\varphi(t)\) is odd, i. e., it holds that \(q := \gcd\{k : \varphi_k > 0\} \equiv 1 \pmod 2\), and finally (iv) we assume that there are at least three different indices with a positive degree-weight, i. e., \(|\{k : \varphi_k > 0\}| \geq 3\).

We want to remark that if the conditions (iii) and (iv) are not satisfied then the limiting distribution of the path edge-covering numbers \(P_n\) degenerates from Gaussian to a point distribution, i. e., the whole mass is concentrated at one point.

The asymptotic behaviour of \(T(z)\) as solution of (2) is discussed in detail in [4]. It follows that \(T(z)\) has a dominant singularity at \(z = \rho\) with \(\rho = \frac{\tau}{\varphi(\tau)} = \frac{1}{\varphi'(\tau)}\), where \(\tau\) is defined as above. Moreover, \(T(z)\) has the following local expansion around \(z = \rho\) (with \(c\) a certain constant):
\[
T(z) = \tau - \sqrt{\frac{2 \varphi(\tau)}{\varphi'(\tau)}} \sqrt{1 - \frac{z}{\rho}} + c(1 - \frac{z}{\rho}) + O\left((1 - \frac{z}{\rho})^{\frac{3}{2}}\right).
\]
It follows further that for aperiodic degree-weight generating functions \(\varphi(t)\), i. e., \(q = 1\) with \(q\) defined as above, \(\rho\) is the only dominant singularity of \(T(z)\).

For a period \(q > 1\) the function \(T(z)\) has \(q\) dominant singularities \(z = q_j, 0 \leq j \leq q - 1\), with \(q_j := \omega^j \rho\) and \(\omega\) a primitive \(q\)-th root of unity. Using the abbreviation \(\tau_j := \omega^j \tau\), one obtains the following local expansion of \(T(z)\) around the singularity \(z = \rho_j\):
\[
T(z) = \tau_j - \omega^j \sqrt{\frac{2 \varphi(\tau)}{\varphi'(\tau)}} \sqrt{1 - \frac{z}{\rho_j}} + c_j(1 - \frac{z}{\rho_j}) + O\left((1 - \frac{z}{\rho_j})^{\frac{3}{2}}\right),
\]
with \(c_j\) certain constants.

2.1.2. Increasing tree families. Increasing trees are labelled trees where the nodes of a tree of size \(n\) are labelled by distinct integers of the set \(\{1, \ldots, n\}\) in such a way that each sequence of labels along any branch starting at the root is increasing. As the underlying tree model we use the simply generated trees but, additionally, they are equipped with increasing labellings. These increasing tree models were introduced in [1].
Given a planted plane tree $T$, we define the weight $w(T)$ as done for simply generated trees (for a given degree-weight sequence $(\varphi_k)_{k \geq 0}$), and denote by $L(T)$ the number of different increasing labellings of the tree $T$. The resulting increasing tree family consists then of all trees $T$ with $w(T) \neq 0$ together with their weights $w(T)$ and the various increasing labellings. It follows then that the exponential generating function $T(z) = \sum_{n \geq 0} T_n z^n$ of the quantity $T_n := \sum_{|T|=n} w(T) \cdot L(T)$ satisfies the following differential equation (where $\varphi(t)$ is defined as for simply generated trees):

$$T'(z) = \varphi(T(z)), \quad T(0) = 0. \quad (4)$$

Three specific increasing tree families are of particular interest and the behaviour of the path edge-covering number is studied for them. Recursive trees are the family of non-plane increasing trees such that all node degrees are allowed. The degree-weight generating function is $\varphi(t) = e^t$. Plane-oriented recursive trees (also called heap ordered trees) are the family of plane increasing trees such that all node degrees are allowed. The degree-weight generating function is $\varphi(t) = \frac{1}{1 - t}$. Binary increasing trees have the degree-weight generating function $\varphi(t) = (1 + t)^2$. This model is also of special importance, since it is isomorphic to the model of binary search trees.

2.1.3. Rooted and unrooted unlabelled trees. Rooted unlabelled trees can be defined recursively as a root followed by a possibly empty set of rooted unlabelled trees, thus a subtree structure and all its permutations are just counted once. If $T_n$ denotes the number of rooted unlabelled trees of size $n$ and $T(z) := \sum_{n \geq 1} T_n z^n$ its ordinary generating function, then $T(z)$ satisfies the functional equation

$$T(z) = z \sum_{k \geq 0} Z(S_k; T(z), T(z^2), \ldots, T(z^k)) = z \sum_{k \geq 0} Z_k\{T(z)\}, \quad (5)$$

where $Z(S_k; x_1, \ldots, x_k)$ denotes the cycle-index of the symmetric group $S_k$ of degree $k$. Throughout this paper, we use for arbitrary functions $f(z)$ the abbreviation

$$Z_k\{f(z)\} := Z(S_k; f(z), f(z^2), \ldots, f(z^k)).$$

Due to the well known identity (see, e.g., [10, p. 590])

$$\sum_{k \geq 0} Z_k\{f(z)\} t^k = \exp \left( \sum_{k \geq 1} \frac{f(z^k) t^k}{k} \right), \quad (6)$$

for any function $f(z)$, this functional equation can be written as

$$T(z) = z \exp \left( \sum_{k \geq 1} \frac{T(z^k)}{k} \right). \quad (7)$$

If we denote by $\tilde{T}_n$ the corresponding number of unrooted unlabelled trees of size $n$ and $\tilde{T}(z) := \sum_{n \geq 1} T_n z^n$ its ordinary generating function, then Otter [10] showed the following
connection between $T(z)$ and $\tilde{T}(z)$:

$$\tilde{T}(z) = T(z) - \frac{1}{2}(T(z)^2 - T(z^2)).$$  \hfill (8)

2.2. Outline of the proof method. In order to obtain our results for the limiting distributions of $P_n$ in the tree families considered, we will use a generating functions approach. This means we translate the recursive description of the tree families and the parameter studied into differential equations resp. functional equations for suitably defined bivariate generating functions $X(z,v)$ and auxiliary functions $Y(z,v)$. If we denote by $X_n$ the random variable that counts the number of nodes $v$ with odd degree $d(v)$ and by $Y_n$ the random variable that counts the number of nodes $v$ with even out-degree $d^+(v)$ in a random tree of size $n$ (thus we have $P_n = \frac{1}{2}X_n$), then these functions are defined as follows:

$$Y(z,v) := \sum_{n\geq 0} \sum_{m\geq 0} P\{Y_n = m\} T_n z^n v^m \text{ resp. } \sum_{n\geq 0} \sum_{m\geq 0} P\{Y_n = m\} T_n \frac{z^n}{n!} v^m, \tag{9a}$$

$$X(z,v) := \sum_{n\geq 0} \sum_{m\geq 0} P\{X_n = m\} T_n z^n v^m \text{ resp. } \sum_{n\geq 0} \sum_{m\geq 0} P\{X_n = m\} T_n \frac{z^n}{n!} v^m. \tag{9b}$$

The ordinary generating functions are used for simply generated trees and unlabelled trees, whereas the exponential generating functions are used for increasing tree families.

Given a tree $T$, which consists of a root and $d$ subtrees $T_1, \ldots, T_d$, we use thereby the following recursive description of $X(T)$ and the auxiliary quantity $Y(T)$, which is defined as the number of nodes $v$ in $T$ with even out-degree $d^+(v)$, in terms of these quantities of the subtrees $T_1, \ldots, T_d$:

$$Y(T) = [d := \text{degree of the root} = \text{even}] + Y(T_1) + \cdots + Y(T_d), \tag{10a}$$

$$X(T) = [d := \text{degree of the root} = \text{odd}] + Y(T_1) + \cdots + Y(T_d). \tag{10b}$$

We use here the Iverson-notation: $[S]$ is 1 if the statement $S$ is true, and 0 if $S$ is false. The recursive description of the tree families together with the recursive description (10) of the parameters considered immediately leads to the following equations for the generating functions $X(z,v)$ and $Y(z,v)$ as defined by (9).

Functional equations for simply generated trees:

$$Y(z,v) = zv\varphi_e(Y(z,v)) + zv\varphi_u(Y(z,v)) = z\varphi(Y(z,v)) + z(v - 1)\varphi_e(Y(z,v)), \tag{11a}$$

$$X(z,v) = z\varphi_e(Y(z,v)) + zv\varphi_u(Y(z,v)) = vz\varphi(Y(z,v)) + z(1 - v)\varphi_e(Y(z,v)). \tag{11b}$$

Differential equations for increasing trees:

$$\frac{\partial}{\partial z} Y(z,v) = v\varphi_e(Y(z,v)) + \varphi_u(Y(z,v)), \tag{12a}$$

$$\frac{\partial}{\partial z} X(z,v) = \varphi_e(Y(z,v)) + v\varphi_u(Y(z,v)). \tag{12b}$$
Functional equations for rooted unlabelled trees:
\[
Y(z, v) = zv \sum_{k \geq 0} Z_{2k} \{ Y(z, v) \} + z \sum_{k \geq 0} Z_{2k+1} \{ Y(z, v) \}, \tag{13a}
\]
\[
X(z, v) = z \sum_{k \geq 0} Z_{2k} \{ Y(z, v) \} + zv \sum_{k \geq 0} Z_{2k+1} \{ Y(z, v) \}. \tag{13b}
\]

We use there the abbreviations
\[
\varphi_e(t) := \sum_{k \geq 0} \varphi_{2k} t^{2k} = \frac{\varphi(t) + \varphi(-t)}{2}, \quad \text{and} \quad \varphi_u(t) := \sum_{k \geq 0} \varphi_{2k+1} t^{2k+1} = \frac{\varphi(t) - \varphi(-t)}{2},
\]
for a degree-weight generating function \( \varphi(t) = \sum_{k \geq 0} \varphi_k t^k \).

In our analysis we are interested in the asymptotic behaviour of the coefficients \([z^n]X(z, v)\) uniformly in a neighbourhood of \(v = 1\). We can establish local expansions of \(X(z, v)\) around their dominant singularities (singularities of smallest modulus) \(z = \rho(v)\), which hold uniformly for \(|v - 1| \leq \eta\) with \(\eta > 0\). Only for two specific tree families, namely recursive trees and binary increasing trees, we are able to obtain explicit formulæ for \(X(z, v)\), which yield directly the required local expansions. For the remaining tree families we will apply the Theorem of Drmota-Lalley-Woods to establish these local expansions from certain functional equations appearing. We state here this theorem as given in [2].

**Theorem 1.** [Drmota-Lalley-Woods] Suppose that \(F(z, y, v) = \sum_{n,k} F_{n,k}(v) z^n y^k\) is an analytic function in \(z, y,\) and \(v\) around 0 such that \(F(0, y, v) \equiv 0\), that \(F(z, 0, v) \neq 0\), and that all coefficients \(F_{n,k}(1)\) of \(F(z, y, 1)\) are real and non-negative. Then the unique solution \(y = y(z, v) = \sum_{n} y_n(v) z^n\) of the functional equation
\[
y = F(z, y, v) \tag{14}
\]
with \(y(0, v) = 0\) is analytic around 0 and has non-negative coefficients \(y_n(1)\) for \(v = 1\).

Furthermore, if the region of convergence of \(F(z, y, v)\) is large enough such that there exist positive solutions \(z = \rho\) and \(y = \tau\) of the system of equations
\[
y = F(z, y, 1),
\]
\[
1 = F_y(z, y, 1),
\]
with \(F_x(\rho, \tau, 1) \neq 0\) and \(F_{yy}(\rho, \tau, 1) \neq 0\) then there exist functions \(\rho(v), g(z, v), h(z, v)\) which are analytic around \(z = \rho, v = 1\) such that \(y(z, v)\) is analytic for \(|z| < \rho\) and \(|v - 1| \leq \epsilon\) (for some \(\epsilon > 0\)) and has a representation of the form
\[
y(z, v) = g(z, v) - h(z, v) \sqrt{1 - \frac{z}{\rho(v)}}, \tag{15}
\]
locally around \(z = \rho, v = 1\). We have \(\tau(v) := g(\rho(v), v) = y(\rho(v), v)\) and
\[
h(\rho(v), v) = \sqrt{\frac{2\rho(v) F_x(\rho(v), \tau(v))}{F_{yy}(\rho(v), \tau(v))}}. \tag{16}
\]
Moreover, (15) provides a local analytic continuation of \( y(z, v) \) for \( \arg(z - \rho(v)) \neq 0 \). If we further assume that \( y_n(1) > 0 \) for \( n \geq n_0 \) then \( z = \rho(v) \) is the only dominant singularity of \( y(z, v) \) for \( |v - 1| \leq \epsilon \).

The local expansions appearing here are amenable for singularity analysis and we will apply the transfer lemmata of Flajolet and Odlyzko [5], which allow us to transfer the local behaviour of a generating function around its dominant singularities to the asymptotic behaviour of its coefficients. The asymptotic expansions obtained for the coefficients \( [z^n]X(z, v) \) give immediately an asymptotic expansion of the moment generating function (= Laplace transform) \( \mathbb{E}(e^{P_n s}) \) of the random variable \( P_n \) considered via

\[
\mathbb{E}(e^{P_n s}) = \sum_{m \geq 0} \mathbb{P}(P_n = m) T_n e^{ms} = \sum_{m \geq 0} \mathbb{P}(X_n = 2m) T_n e^{ms} = \frac{[z^n]X(z, e^{\frac{s}{2}})}{[z^n]X(z, 1)}. \tag{17}
\]

Then we can apply the continuity theorem of the Laplace transform to obtain the convergence in distribution of \( P_n \) to a Gaussian distributed random variable \( P \). In the instances appearing here we can apply directly the so called quasi-power theorem due to Hwang [7]. It gives a powerful method not only to prove the Gaussian limit law but also to determine the rate of convergence. It is stated below, where \( \Phi(x) \) denotes the distribution function of the standard normal distribution.

**Theorem 2.** [H. K. Hwang] Let \( \{X_n\}_{n \geq 1} \) be a sequence of integral random variables. Suppose that the moment generating function satisfies the asymptotic expression

\[
M_n(s) := \mathbb{E}(e^{X_n s}) = \sum_{m \geq 0} \mathbb{P}(X_n = m) e^{ms} = e^{H_n(s)}(1 + O(\kappa_n^{-1})),
\]

the \( O \)-term being uniform for \( |s| \leq \sigma, \ s \in \mathbb{C}, \ \sigma > 0 \), where

(i) \( H_n(s) = U(s)\phi(n) + V(s) \), with \( U(s) \) and \( V(s) \) analytic for \( |s| \leq \sigma \) and independent of \( n \); \( U''(0) \neq 0 \),

(ii) \( \phi(n) \to \infty \),

(iii) \( \kappa_n \to \infty \).

Under these assumptions, the distribution of \( X_n \) is asymptotically Gaussian with the given convergence rate in the Kolmogorov metric:

\[
\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \leq x \right) - \Phi(x) \right| = O\left( \frac{1}{\kappa_n} + \frac{1}{\sqrt{\phi(n)}} \right).
\]

Moreover, the mean and the variance of \( X_n \) satisfy

\[
\mathbb{E}(X_n) = U'(0)\phi(n) + V'(0) + O(\kappa_n^{-1}), \quad \mathbb{V}(X_n) = U''(0)\phi(n) + V''(0) + O(\kappa_n^{-1}).
\]

The results of this paper are presented in Section 3, and the corresponding proofs are given in Section 4–9.
3. Results for the tree families considered

Theorem 3. Let the tree family $T$ be one of the following: (i) simply generated trees with the restrictions to the degree-weight generating function $\varphi(t)$ as given in Subsection 2.1.1, (ii) recursive trees, (iii) binary increasing trees and binary search trees, (iv) plane-oriented recursive trees, (v) rooted (resp. unrooted) unlabelled trees. Then the distribution of the random variable $P_n$, which counts the path edge-covering number of a random tree of size $n$ in the family $T$, is asymptotically Gaussian, where the convergence rate is of order $O\left(\frac{1}{\sqrt{n}}\right)$:

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left\{ \frac{P_n - \mathbb{E}(P_n)}{\sqrt{\mathbb{V}(P_n)}} \leq x \right\} - \Phi(x) \right| = O\left(\frac{1}{\sqrt{n}}\right).$$

The expectation $\mathbb{E}(P_n)$ and the variance $\mathbb{V}(P_n)$ satisfy the following asymptotic expansions, where $c$ and $d$ are some constants, which can be different on every occurrence.

(i) Simply generated trees:

$$\mathbb{E}(P_n) = \frac{\varphi_e(\tau)}{2\varphi(\tau)} n + c + O(n^{-1}),$$
$$\mathbb{V}(P_n) = \frac{\tau^2 \varphi''(\tau) \varphi_e(\tau) (\varphi(\tau) - \varphi_e(\tau)) - \varphi(\tau) (\varphi_e(\tau) - \tau \varphi'(\tau))^2}{4\varphi''(\tau) \tau^2 \varphi(\tau)^2} n + d + O(n^{-1}).$$

(ii) Recursive trees:

$$\mathbb{E}(P_n) = \frac{1}{3} n, \text{ for } n \geq 3, \quad \mathbb{V}(P_n) = \frac{2}{45} n, \text{ for } n \geq 5.$$

(iii) Binary increasing trees (and binary search trees):

$$\mathbb{E}(P_n) = \frac{1}{3} n - \frac{2}{3} + \frac{2}{n}, \text{ for } n \geq 2, \quad \mathbb{V}(P_n) = \frac{2}{45} n + \frac{2}{45} - \frac{2}{3n} - \frac{4}{n^2}, \text{ for } n \geq 4.$$

(iv) Plane-oriented recursive trees:

$$\mathbb{E}(P_n) = (2 \log(2) - 1)n + \frac{5}{8} - \log(2) + O(n^{-1}),$$
$$\mathbb{V}(P_n) = \frac{4 \log^2(2) - 2 \log(2) - \frac{1}{2}}{32} n + \frac{9}{32} + \log(2) - 2 \log^2(2) + O(n^{-1}).$$

(v) Rooted (resp. unrooted) unlabelled trees:

$$\mathbb{E}(P_n) = \mu n + c + O(n^{-1}), \quad \mathbb{V}(P_n) = \nu n + d + O(n^{-1}),$$

with certain constants $\mu$ and $\nu$. These constants can be specified via solutions of certain functional equations. An expression for $\mu$ is given by equation (81) with $\mu = U'(0)$.

Note that more general results of the path edge-covering number for increasing tree families are available. In particular, it was shown by the author that $P_n$ is for all polynomial increasing tree families, i.e., tree families where the node-degrees of the trees are bounded by some number $d$ or equivalently the degree-weight generating function $\varphi(t)$ are polynomials, asymptotically Gaussian distributed. Since the proof of this result is shown via the method
of moments, which leads to more lengthy computations, it is not included in the present paper.

4. Simply generated trees

We will carry out here the proof of Theorem 3 only for aperiodic degree-weight generating functions \( \varphi(t) \), but mention that the proof is fully analogous for \( \varphi(t) \) with a period \( q > 1 \), where \( q \) is odd due to the restriction (iii) given in Subsection 2.1.1. Then one has to take into account the contributions of all \( q \) dominant singularities in the corresponding expansions of the generating functions appearing.

First we will apply Theorem 1 to equation (11a), i.e., we study

\[
Y = F(z, Y, v) := z \varphi(Y) + z(v - 1)\varphi_e(Y). \tag{18}
\]

It is seen easily that the assumptions on \( F(z, Y, v) \) made in Theorem 1 are satisfied: we have \( F(0, Y, v) \equiv 0, F(z, 0, v) = \varphi_0 z v \neq 0 \), and the system \( Y = F(z, Y, 1), 1 = F_Y(z, Y, 1) \) is given by

\[
Y = z \varphi(Y), \quad 1 = z \varphi'(Y), \tag{19}
\]

which leads to the equations \( Y \varphi'(Y) = \varphi(Y) \) and \( z = \frac{Y}{\varphi(Y)} = \frac{1}{\varphi'(Y)} \). But due to the restrictions (i) and (ii) made on the degree-weight generating functions \( \varphi(t) \) in Subsection 2.1.1 it is guaranteed that there exists a minimal positive solution \( Y = \tau \) of the equation \( Y \varphi'(Y) = \varphi(Y) \), which also leads to a positive solution for \( z: z = \rho = \frac{\tau}{\varphi(\tau)} \). Thus the solution \( z = \rho \) and \( Y = \tau \) of the system (19) leads exactly to the constants described in Subsection 2.1.1. It further holds that \( F_z(\rho, \tau, 1) = \varphi(\tau) > 0 \) and \( F_{YY}(\rho, \tau, 1) = \rho \varphi''(\tau) > 0 \). Therefore \( Y(z, v) \) has the representation

\[
Y(z, v) = g(z, v) - h(z, v) \sqrt{1 - \frac{z}{\rho(v)}}, \tag{20}
\]

with functions \( \rho(v), g(z, v), \) and \( h(z, v) \) that are analytic around \( z = \rho \) and \( v = 1 \).

Furthermore, using (16) we obtain from (18) the following local expansion around the unique dominant singularity \( z = \rho(v) \) (with \( c(v) \) a certain function; of course it holds \( \rho(1) = \rho \) and \( \tau(1) = \tau \):

\[
Y(z, v) = \tau(v) - \sqrt{\frac{2(\varphi(\tau(v)) + (v - 1)\varphi_e(\tau(v)))}{\varphi''(\tau(v)) + (v - 1)\varphi''_e(\tau(v))}} \sqrt{1 - \frac{z}{\rho(v)}}
+ c(v)(1 - \frac{z}{\rho(v)}) + O\left((1 - \frac{z}{\rho(v)})^{\frac{3}{2}}\right). \tag{21}
\]

Next we consider the function \( X(z, v) \) as given by equation (11b). We obtain from the representation (20) of \( Y(z, v) \) that \( X(z, v) \) has a corresponding representation with analytic functions \( g(z, v) \) and \( h(z, v) \) around the same unique dominant singularity \( z = \rho(v) \). To obtain a local expansion of \( X(z, v) \) around \( z = \rho(v) \) in a neighbourhood of \( v = 1 \) we use
(21) and apply in (11b) a Taylor series expansion of $\varphi(t)$ resp. $\varphi_e(t)$ around $t = \tau(v)$ (with $\hat{c}(v)$ a certain function):

\[
X(z, v) = (\rho(v) + (z - \rho(v))) \times \\
\times \left[ v \left( \varphi(\tau(v)) + \varphi'(\tau(v))(Y(z, v) - \tau(v)) + \frac{\varphi''(\tau(v))}{2} (Y(z, v) - \tau(v))^2 + \mathcal{O}((Y(z, v) - \tau(v))^3) \right) \\
+ (1 - v) \left( \varphi_e(\tau(v)) + \varphi_e'(\tau(v))(Y(z, v) - \tau(v)) + \frac{\varphi''_e(\tau(v))}{2} (Y(z, v) - \tau(v))^2 \\
+ \mathcal{O}((Y(z, v) - \tau(v))^3) \right) \right] \\
= \rho(v) (v \varphi(\tau(v)) + (1 - v) \varphi_e(\tau(v))) \\
- \rho(v) (v \varphi'(\tau(v)) + (1 - v) \varphi_e'(\tau(v))) \sqrt{\frac{2(\varphi(\tau(v)) + (v - 1) \varphi_e(\tau(v)))}{\varphi''(\tau(v)) + (v - 1) \varphi''_e(\tau(v))}} \sqrt{1 - \frac{z}{\rho(v)}} \tag{22} \\
+ \hat{c}(v) \left( 1 - \frac{z}{\rho(v)} \right) + \mathcal{O} \left( \left( 1 - \frac{z}{\rho(v)} \right)^{\frac{3}{2}} \right).
\]

Singularity analysis [5, Proposition 1 and Theorem 1] leads thus from (22) immediately to the following asymptotic expansion of the coefficients:

\[
[z^n]X(z, v) = \frac{\rho(v)(v \varphi'(\tau(v)) + (1 - v) \varphi_e'(\tau(v)))}{2\sqrt{\pi}} \sqrt{\frac{2(\varphi(\tau(v)) + (v - 1) \varphi_e(\tau(v)))}{\varphi''(\tau(v)) + (v - 1) \varphi''_e(\tau(v))}} \times \tag{23}
\times (\rho(v))^{-n} n^{-\frac{3}{2}} \left( 1 + \mathcal{O}(n^{-1}) \right).
\]

Together with the asymptotic expansion

\[
[z^n]X(z, 1) = [z^n]T(z) = \frac{\rho^{-n} n^{-\frac{3}{2}}}{2 \sqrt{\pi}} \sqrt{\frac{2 \varphi(\tau)}{\varphi''(\tau)}} \left( 1 + \mathcal{O}(n^{-1}) \right), \tag{24}
\]

which again is obtained via singularity analysis from (3), we obtain from (23) and (24) by using (17) the following expansion of the moment generating function:

\[
\mathbb{E}(e^{P_n s}) = \exp \left( n U(s) + V(s) \right) \left( 1 + \mathcal{O}(n^{-1}) \right), \tag{25}
\]

where

\[
U(s) = \log \left( \frac{\rho}{\rho(e^{s})} \right), \tag{26}
\]

and

\[
V(s) = \log \left( \rho(e^{s}) \right) + \log \left( e^{s} \varphi'(e^{s}) \right) + (1 - e^{s}) \varphi_e'(e^{s}) \\
+ \frac{1}{2} \log \left( \frac{\varphi''(e^{s})}{\varphi(e^{s})} \left( \varphi'(e^{s}) \right) + (e^{s} - 1) \varphi_e(e^{s}) \right) .
\]

The corresponding part of Theorem 3 concerning the Gaussian limiting distribution follows now by an application of Theorem 2, provided that $U''(0) \neq 0$ holds (see remarks below).

In order to obtain the leading constants $U'(0)$ and $U''(0)$ in the expansion of the expectation $\mathbb{E}(P_n)$ resp. the variance $\mathbb{V}(P_n)$ we have to study the behaviour of the functions $\rho(v)$
(and $\tau(v)$) around $v = 1$. Plugging in $z = \rho(v)$ and $Y = \tau(v)$ to the system of equations
$Y = F(z, Y, v), \quad 1 = F_Y(z, Y, v)$ gives then
\begin{equation}
\tau(v) = \rho(v)(\varphi(\tau(v)) + (v - 1)\varphi_e(\tau(v))),
\end{equation}
\begin{equation}
1 = \rho(v)(\varphi'(\tau(v)) + (v - 1)\varphi'_e(\tau(v))).
\end{equation}

Since we obtain from (26)
\begin{equation}
U'(0) = -\frac{\rho'(1)}{2\rho}, \quad \text{and} \quad U''(0) = \frac{(\rho'(1))^2 - \rho\rho'(1) - \rho\rho''(1)}{4\rho^2},
\end{equation}
we have to determine $\rho'(1)$ and $\rho''(1)$.

Differentiating (27a) w. r. t. $v$ and evaluating at $v = 1$ gives
\begin{equation}
\tau'(1) = \rho'(1)\varphi(\tau) + \rho(\varphi'(\tau))\tau'(1) + \varphi_e(\tau),
\end{equation}
which implies
\begin{equation}
\tau'(1)(1 - \rho\varphi'(\tau)) = \rho'(1)\varphi(\tau) + \rho\varphi_e(\tau).
\end{equation}
Due to $1 - \rho\varphi'(\tau) = 0$, we further obtain
\begin{equation}
\rho'(1) = -\frac{\rho\varphi_e(\tau)}{\varphi(\tau)}. \quad (29)
\end{equation}
Using (28) and (29) we get thus
\begin{equation}
U'(0) = \frac{\varphi_e(\tau)}{2\varphi(\tau)} = \frac{1}{4} \left( 1 + \frac{\varphi(-\tau)}{\varphi(\tau)} \right),
\end{equation}
which was already obtained in [8].

To get also $\rho''(1)$ we require $\tau'(1)$. This quantity can be obtained from (27b) after differentiating w. r. t. $v$ and evaluating at $v = 1$ by using (29). We get
\begin{equation}
\tau'(1) = \frac{\varphi_e(\tau)\varphi'(\tau) - \varphi(\tau)\varphi'_e(\tau)}{\varphi(\tau)\varphi''(\tau)}. \quad (31)
\end{equation}
Now differentiating (27a) twice w. r. t. $v$ and evaluating at $v = 1$ leads to
\begin{equation}
\rho''(1) = -\frac{2\rho'(1)\varphi_e(\tau) + 2\rho'(1)\varphi'(\tau)\tau'(1) + 2\rho\varphi'_e(\tau)\tau'(1) + \rho\varphi''(\tau)(\tau'(1))^2}{\varphi(\tau)}. \quad (32)
\end{equation}
Combining the results (28), (29), (31) and (32) gives then:
\begin{equation}
U''(0) = \frac{\tau^2\varphi''(\tau)\varphi_e(\tau)(\varphi(\tau) - \varphi_e(\tau)) - \varphi(\tau)(\varphi_e(\tau) - \tau\varphi'_e(\tau))^2}{4\varphi''(\tau)\tau^2\varphi(\tau)^2}. \quad (33)
\end{equation}

In order to show that $P_n$ is actually asymptotically Gaussian it remains to prove that $U''(0) > 0$. But obviously there are instances of degree-weight generating functions $\varphi(t)$, where $U''(0) = 0$, i. e., the limiting distribution degenerates. E. g., this is the case if the period $q$ is even: then all nodes (apart from the root) in an arbitrary size-$n$ tree of such a tree family have odd degree and therefore give a contribution to $X_n$ resp. $P_n$. 

In order to characterize all suitable \( \varphi(t) \) as done here via the assumptions (iii) and (iv), one can consider the numerator of \( U''(0) \) in (33):

\[
N(\tau) := \tau^2 \varphi''(\tau) \varphi_e(\tau) (\varphi(\tau) - \varphi_e(\tau)) - \varphi(\tau)(\varphi_e(\tau) - \tau \varphi'_e(\tau))^2,
\]

(34)
carrying out a power series expansion: \( N(\tau) := \sum_{k \geq 0} N_k \tau^k \), and checking the existence of positive coefficients \( N_k > 0 \).

But when doing this one has to take into account the relation \( \varphi(\tau) = \tau \varphi'(\tau) \), which gives

\[
\varphi_0 = \sum_{k \geq 1} (k - 1) \varphi_k \tau^k.
\]

(35)

This leads to the expansions

\[
\varphi(\tau) = \sum_{k \geq 1} k \varphi_k \tau^k, \quad \varphi_e(\tau) = \sum_{k \geq 1} 2k \varphi_{2k} \tau^{2k} + \sum_{k \geq 1} 2k \varphi_{2k+1} \tau^{2k+1},
\]

\[
\varphi_e(\tau) - \tau \varphi'_e(\tau) = \sum_{k \geq 1} 2k \varphi_{2k+1} \tau^{2k+1}, \quad \tau^2 \varphi''(\tau) = \sum_{k \geq 1} k(k-1) \varphi_k \tau^k,
\]

\[
\varphi(\tau) - \varphi_e(\tau) = \sum_{k \geq 1} \varphi_{2k-1} \tau^{2k-1}.
\]

Plugging in these expansions into (34) it can be shown that the assumptions (iii) and (iv) made on \( \varphi(t) \) are necessary and sufficient for \( U''(0) > 0 \). These considerations are here omitted, since they are, although elementary, a bit lengthy.

5. Recursive trees

Specifying \( \varphi(t) = e^t \) in equations (12) gives the following system of differential equations for the generating functions \( X(z, v) \) and \( Y(z, v) \):

\[
\frac{\partial}{\partial z} Y(z, v) = \frac{v e^{Y(z, v)} + e^{-Y(z, v)}}{2} + \frac{e^{Y(z, v)} - e^{-Y(z, v)}}{2}, \quad (36a)
\]

\[
\frac{\partial}{\partial z} X(z, v) = \frac{e^{Y(z, v)} + e^{-Y(z, v)}}{2} + v \frac{e^{Y(z, v)} - e^{-Y(z, v)}}{2}. \quad (36b)
\]

We easily get from (36a) by separating variables the solution of \( Y(z, v) \) implicitly via

\[
z = \int_0^{Y(z, v)} \frac{2 dt}{(v + 1)e^t + (v - 1)e^{-t}}
\]

\[
= \frac{1}{\sqrt{1 - v^2}} \log \left( \frac{\sqrt{\frac{v+1}{1-v}}e^{Y(z, v)} - 1}{\sqrt{\frac{v+1}{1-v}}e^{Y(z, v)} + 1} \right) - \frac{1}{\sqrt{1 - v^2}} \log \left( \frac{\sqrt{\frac{v+1}{1-v}} - 1}{\sqrt{\frac{v+1}{1-v}} + 1} \right),
\]

which leads to the following explicit solution:

\[
Y(z, v) = \log \left( \frac{\sqrt{\frac{v+1}{1+v}} \left( \sqrt{\frac{v+1}{1-v}} - 1 \right) e^{\sqrt{1-v^2}z} + 1}{\sqrt{\frac{v+1}{1+v}} - 1 \left( \sqrt{\frac{v+1}{1-v}} + 1 \right) e^{\sqrt{1-v^2}z}} \right). \quad (37)
\]
We get then that in a neighbourhood of $v = 1$ the dominant singularity $z = \rho(v)$ of $Y(z, v)$ is given by
\[
\rho(v) = \frac{1}{\sqrt{1 - v^2}} \log \left( \frac{\sqrt{v + 1} + \sqrt{1 - v}}{\sqrt{v + 1} - \sqrt{1 - v}} \right),
\]
where the denominator of the argument of the logarithm in (37) vanishes. It is seen easily either by expanding $\rho(v)$ or by the integral representation $\rho(v) = \int_{0}^{\infty} \frac{2dt}{(v+1)e^{t}+(v-1)e^{-t}}$ that $\rho(v)$ is analytic in a neighbourhood of $v = 1$ and it holds $\rho(1) = 1$.

Expanding $Y(z, v)$ as given by (37) around the dominant singularity $z = \rho(v)$ is an easy task and leads to the following local expansion, which holds uniformly around $v = 1$:
\[
Y(z, v) = \log \left( \frac{1}{1 - \frac{z}{\rho(v)}} \right) + \log \left( \frac{2}{(v+1)\rho(v)} \right) + O(1 - \frac{z}{\rho(v)}),
\]
(39)

Next we study $X(z, v)$: from the differential equation (36b) we immediately obtain the solution
\[
X(z, v) = \frac{1+v}{2} \int_{0}^{z} e^{Y(t,v)} dt + \frac{1-v}{2} \int_{0}^{z} e^{-Y(t,v)} dt.
\]
(40)
Using the expansion (39) for $Y(z, v)$ gives
\[
e^{Y(t,v)} = \frac{2}{(v+1)\rho(v)(1 - \frac{t}{\rho(v)})} \left( 1 + O\left(1 - \frac{t}{\rho(v)} \right) \right),
\]
e^{-Y(t,v)} = \frac{(v+1)\rho(v)}{2} \left( 1 - \frac{t}{\rho(v)} \right) \left( 1 + O\left(1 - \frac{t}{\rho(v)} \right) \right).

Applying a theorem on the integration of asymptotic expansions (see, e. g., [3]) we further get
\[
\int_{0}^{z} e^{Y(t,v)} dt = \frac{2}{v+1} \log \left( \frac{1}{1 - \frac{z}{\rho(v)}} \right) + c_{1}(v) + O\left(1 - \frac{z}{\rho(v)} \right),
\]
(41a)
\[
\int_{0}^{z} e^{-Y(t,v)} dt = c_{2}(v) + O\left(\left(1 - \frac{z}{\rho(v)} \right)^{2} \right),
\]
(41b)
with certain functions $c_{1}(v), c_{2}(v)$. Thus we obtain by combining (40) and (41) the required local expansion of $X(z, v)$ around the dominant singularity $z = \rho(v)$, which holds uniformly in a neighbourhood of $v = 1$ (with $c(v)$ a certain function):
\[
X(z, v) = \log \left( \frac{1}{1 - \frac{z}{\rho(v)}} \right) + c(v) + O\left(1 - \frac{z}{\rho(v)} \right).
\]
(42)

Applying singularity analysis [5, Theorem 1] to (42) leads thus to the asymptotic expansion
\[
[z^{n}]X(z, v) = \frac{1}{n} \rho(v)^{-n} \left( 1 + O(n^{-1}) \right).
\]
(43)

Using equation (43) and $[z^{n}]X(z, 1) = [z^{n}]T(z) = [z^{n}]\log \frac{1}{1 - z} = \frac{1}{n}$ we get from (17) the required expansion of the moment generating function:
\[
E(e^{P_{n}s}) = \rho(e^{\frac{s}{n}})^{-n} \left( 1 + O(n^{-1}) \right) = \exp \left( n \log \left( \rho(e^{\frac{s}{n}})^{-1} \right) \right) \left( 1 + O(n^{-1}) \right).
\]
(44)
Thus we can apply Theorem 2 with $U(s) = \log (\rho(e^{s})^{-1})$ and $V(s) = 0$. Differentiating $U(s)$ w. r. t. $s$ and evaluating at $s = 0$ shows then asymptotic normality of the parameter studied with the following expansions of the expectation resp. the variance:

$$E(P_n) = \frac{1}{3} n + O(n^{-1}), \quad \text{Var}(P_n) = \frac{2}{45} n + O(n^{-1}).$$

The refined results for $E(P_n)$ and $\text{Var}(P_n)$ as given in Theorem 3 are already computed by Meir and Moon [8] via different considerations, but it is possible to obtain them from the explicit solution (37).

6. Binary increasing trees and binary search trees

Specifying $\varphi(t) = (1 + t)^2$ in equations (12) gives the system of differential equations

$$\frac{\partial}{\partial z} Y(z, v) = vY(z, v)^2 + 2Y(z, v) + v, \quad (45a)$$

$$\frac{\partial}{\partial z} X(z, v) = Y(z, v)^2 + 2vY(z, v) + 1. \quad (45b)$$

We obtain the following explicit solution of (45a):

$$Y(z, v) = \frac{v \tan(z\sqrt{v^2 - 1})}{\sqrt{v^2 - 1} - \tan(z\sqrt{v^2 - 1})}. \quad (46)$$

The dominant singularity $z = \rho(v)$ of $Y(z, v)$ is in a neighbourhood of $v = 1$ given by

$$\rho(v) = \frac{1}{\sqrt{v^2 - 1}} \arctan \left( \sqrt{v^2 - 1} \right), \quad (47)$$

since here the denominator of $Y(z, v)$ as given by (46) vanishes. Expanding $Y(z, v)$ around $z = \rho(v)$ leads then to the local expansion (with $c(v)$ a certain function)

$$Y(z, v) = \frac{1}{vp(v)(1 - z/\rho(v))} + c(v) + O\left(1 - \frac{z}{\rho(v)}\right), \quad (48)$$

which holds uniformly in a neighbourhood of $v = 1$.

Plugging in expansion (48) into the obvious solution of the differential equation (45b) we obtain also the following local expansion of $X(z, v)$ around the dominant singularity $z = \rho(v)$, which again holds uniformly around $v = 1$ (with certain functions $c_1(v)$ and $c_2(v)$):

$$X(z, v) = \frac{1}{v^2\rho(v)(1 - z/\rho(v))} + c_1(v) \log \left( \frac{1}{1 - z/\rho(v)} \right) + c_2(v) + O\left(1 - \frac{z}{\rho(v)}\right). \quad (49)$$

Singularity analysis [5, Theorem 1] leads then from (49) immediately to the following asymptotic expansion of the coefficients:

$$[z^n]X(z, v) = \frac{1}{v^2\rho(v)n+1}(1 + O(n^{-1})). \quad (50)$$
Since \([z^n]X(z,1) = [z^n]T(z) = [z^n]\frac{1}{1-z} = 1\), equation (50) leads via (17) to the following expansion of the moment generating function:

\[
E(e^{P_n s}) = \exp \left( n \log \left( \rho(e^{\frac{s^2}{2}}) \right) - s - \log \left( \rho(e^{\frac{s^2}{2}}) \right) \right) (1 + O(n^{-1})). \quad (51)
\]

Thus we can apply Theorem 2 with \(U(s) = \log \left( \rho(e^{\frac{s^2}{2}}) \right)\) and \(V(s) = -s - \log \left( \rho(e^{\frac{s^2}{2}}) \right)\). Differentiating \(U(s)\) and \(V(s)\) w. r. t. \(s\) and evaluating at \(s = 0\) shows then asymptotic normality of the parameter studied. The refined results for \(E(P_n)\) and \(V(P_n)\) as given in Theorem 3 are obtained from the explicit solution (46).

7. Plane-oriented recursive trees

Here we obtain from (12) with \(\varphi(t) = \frac{1}{1-t}\) the system of differential equations

\[
\frac{\partial}{\partial z} Y(z, v) = \frac{v + Y(z, v)}{1 - Y(z, v)^2}, \quad (52a)
\]

\[
\frac{\partial}{\partial z} X(z, v) = \frac{1 + vY(z, v)}{1 - Y(z, v)^2}. \quad (52b)
\]

The solution of equation (52a) is given implicitly by

\[
z = \int_0^{Y(z,v)} \frac{1 - t^2}{v + t} dt = vY(z, v) - \frac{1}{2}Y(z, v)^2 + (1-v^2) \log(v + Y(z, v)) - (1-v^2) \log v. \quad (53)
\]

To obtain the behaviour of the generating function \(Y(z, v)\) around \(v = 1\) we will apply Theorem 1, where we first rewrite equation (53):

\[
Y = F(z, Y, v) := \frac{1}{2}Y^2 + (1-v)Y - (1-v^2) \log(v+Y) + (1-v^2) \log v + z. \quad (54)
\]

It is then seen easily that the assumptions on \(F(z, Y, v)\) made in Theorem 1 are satisfied and that \(\tau = 1\) and \(\rho = \frac{1}{2}\). Therefore \(Y(z, v)\) has the representation

\[
Y(z, v) = g(z, v) - h(z, v) \sqrt{1 - \frac{z}{\rho(v)}}, \quad (55)
\]

with functions \(\rho(v), g(z, v), h(z, v)\) that are analytic around \(z = \rho\) and \(v = 1\). To give \(\rho(v)\) and \(\tau(v) = Y(\rho(v), v) = g(\rho(v), v)\) explicitly we plug in \(z = \rho(v)\) and \(Y = \tau(v)\) to the system of functional equations \(Y = F(z, Y, v)\), \(1 = F_Y(z, Y, v)\). From the equation

\[
1 = F_Y(z, Y, v) = Y + 1 - v - \frac{1 - v^2}{v + Y}
\]

follows that \(Y^2 = 1\) and thus that \(\tau(v) = 1\) in a neighbourhood of \(v = 1\). With this result we obtain from (54) also an explicit expression for the dominant singularity \(z = \rho(v)\):

\[
\rho(v) = v - \frac{1}{2} + (1-v^2) \log(v+1) - (1-v^2) \log v. \quad (56)
\]

Using (16) yields then the following local expansion of \(Y(z, v)\) around \(z = \rho(v)\) (with \(c(v)\) a certain function):

\[
Y(z, v) = 1 - \sqrt{1 + v\sqrt{\rho(v)}} \sqrt{1 - \frac{z}{\rho(v)}} + c(v)(1 - \frac{z}{\rho(v)}) + O\left((1 - \frac{z}{\rho(v)})^\frac{3}{2}\right), \quad (57)
\]
which holds uniformly in a neighbourhood of \( v = 1 \).

From the solution

\[
X(z, v) = \int_0^z \frac{1 + vY(t, v)}{1 - Y(t, v)}^2 dt
\]

(58)
of (52b) we obtain by using the expansion (57) also an expansion of \( X(z, v) \) around the dominant singularity \( z = \rho(v) \) (with certain functions \( c_0(v), \hat{c}(v) \) and \( \hat{c}(v) \)):

\[
X(z, v) = \int_0^{\rho(v)} \frac{1 + vY(t, v)}{1 - Y(t, v)^2} dt + \int_0^z \frac{1 + vY(t, v)}{1 - Y(t, v)^2} dt
\]

\[
= c_0(v) + \frac{\sqrt{1 + v}}{2\sqrt{\rho(v)}} \int_{\rho(v)}^z \left\{ \frac{1}{\sqrt{1 - \frac{t}{\rho(v)}}} + \hat{c}(v) + \mathcal{O}\left(\frac{1}{\rho(v)}\right) \right\} dt
\]

\[
= c_0(v) - \sqrt{1 + v}\sqrt{\rho(v)} \left\{ \frac{z}{\rho(v)} + \hat{c}(v)(1 - \frac{z}{\rho(v)}) + \mathcal{O}\left(\frac{z}{\rho(v)}\right) \right\},
\]

(59)

which holds uniformly in a neighbourhood of \( v = 1 \). Again we have used here a theorem on the integration of asymptotic expansions that was mentioned previously.

Singularity analysis [5, Proposition 1 and Theorem 1] leads then from (59) to the following asymptotic expansion of the coefficients:

\[
[z^n]X(z, v) = \frac{\sqrt{1 + v}}{2\sqrt{\pi}} \rho(v)^{-n + \frac{1}{2}} n^{-\frac{3}{2}} (1 + \mathcal{O}(n^{-1})).
\]

(60)

Together with

\[
[z^n]X(z, 1) = \frac{2n^{-\frac{3}{2}}}{2\sqrt{\pi}} (1 + \mathcal{O}(n^{-1})),
\]

(61)

which follows from \( X(z, 1) = T(z) = 1 - \sqrt{1 - 2e} \) via singularity analysis, we obtain from (17) the following expansion of the moment generating function:

\[
E(e^{P_n s}) = \exp \left( n \log ((2 \rho(e^{\frac{z}{2}}))^{-1}) + \log \left( \sqrt{1 + e^{\frac{z}{2}}} \sqrt{\rho(e^{\frac{z}{2}})} \right) \right) (1 + \mathcal{O}(n^{-1})).
\]

(62)

With the notation of Theorem 2 we obtain

\[
U(s) = -\log(2\rho(e^{\frac{z}{2}})), \quad \text{and} \quad V(s) = \frac{1}{2} \log(1 + e^{\frac{z}{2}}) + \frac{1}{2} \log(\rho(e^{\frac{z}{2}})),
\]

(63)

and differentiating \( U(s) \) and \( V(s) \) w. r. t. \( s \) and evaluating at \( s = 0 \) shows then the asymptotic normality of the parameter studied with the expansions given in Theorem 3.

### 8. Rooted unlabelled trees

We obtain by using identity (6) the following system of functional equations equivalent to (13):

\[
Y(z, v) = \frac{z(v + 1)}{2} \exp \left( \sum_{k \geq 1} \frac{Y(z, v^k)}{k} \right) + \frac{z(v - 1)}{2} \exp \left( \sum_{k \geq 1} \frac{Y(z, v^k)}{k} (-1)^k \right),
\]

(64a)

\[
X(z, v) = \frac{z(v + 1)}{2} \exp \left( \sum_{k \geq 1} \frac{Y(z, v^k)}{k} \right) + \frac{z(1 - v)}{2} \exp \left( \sum_{k \geq 1} \frac{Y(z, v^k)}{k} (-1)^k \right).
\]

(64b)
We will apply Theorem 1 to (64a) studying the functional equation

\[ Y = F(Y, z, v) := \frac{z(v+1)}{2} e^Y \exp \left( \sum_{k \geq 2} \frac{Y(z^k, v^k)}{k} \right) + \frac{z(v-1)}{2} e^{-Y} \exp \left( \sum_{k \geq 2} \frac{Y(z^k, v^k)}{k} (-1)^k \right). \]

(65)

In order to apply Theorem 1 one has to ensure that the functions \( Y(z^k, v^k) \) are for all \( k \geq 2 \) analytic in a neighbourhood of \( z = \rho \) and \( v = 1 \). But this follows easily due to majorization arguments via

\[ |Y(z^k, v^k)| \leq T(\rho^k) \leq T(\rho), \]

for \( |z| \leq \rho + \eta \) and \( |v| \leq \frac{\sqrt{\eta}}{\rho + \eta} \), where \( \eta > 0 \) is chosen small enough such that \( \frac{\sqrt{\eta}}{\rho + \eta} > 1 \) (see [2, Theorem 14]).

Thus we obtain the following representation of \( Y(z, v) \) in a neighbourhood of the unique dominant singularity \( z = \rho(v) \) around \( v = 1 \):

\[ Y(z, v) = g(z, v) - h(z, v) \sqrt{1 - \frac{z}{\rho(v)}}, \]

(66)

with certain functions \( g(z, v) \) and \( h(z, v) \) analytic around \( z = \rho(v) \) and \( v = 1 \). In particular we get from (16) the local expansion (with \( c(v) \) a certain function)

\[ Y(z, v) = \tau(v) - \kappa(v) \sqrt{1 - \frac{z}{\rho(v)} + c(v) \left(1 - \frac{z}{\rho(v)} \right) + O \left( \left(1 - \frac{z}{\rho(v)} \right)^{\frac{3}{2}} \right)}, \]

(67)

where \( \kappa(v) := \sqrt{\frac{2\rho(v) F_{\tau(p(v), \tau(v))}}{F_{Y, Y(p(v), \tau(v))}}} \). Further it holds \( \tau(1) = \tau = 1 \), \( \rho(1) = \rho \approx 0.338219 \), and \( \kappa(1) =: \kappa \approx 1.559252 \) (see [2]).

From (66) and (67) we obtain easily the required information on the behaviour of the function \( X(z, v) \). Adding the equations (64) leads to

\[ X(z, v) = z(v+1) e^{Y(z, v)} \exp \left( \sum_{k \geq 2} \frac{Y(z^k, v^k)}{k} \right) - Y(z, v). \]

(68)

Since the functions \( Y(z^k, v^k) \) are for \( k \geq 2 \) analytic if \( z \) is close to \( \rho \) and \( v \) close to 1 (see previous remarks), we obtain that \( X(z, v) \) has its unique dominant singularity at \( z = \rho(v) \) with the representation

\[ X(z, v) = \hat{g}(z, v) - \hat{h}(z, v) \sqrt{1 - \frac{z}{\rho(v)}}, \]

(69)

with functions \( \hat{g}(z, v) \) and \( \hat{h}(z, v) \) that are analytic around \( z = \rho(v) \) and \( v = 1 \). Moreover, we obtain by using the expansion (67) the following local expansion of \( X(z, v) \) around \( z = \rho(v) \)
X(z, v) = \rho(v)(v + 1)e^{\tau(v)} \exp \left( \sum_{k \geq 2} \frac{Y(\rho(v)^k, v^k)}{k} \right) - \tau(v) - 1 = 0. \quad (72)

which follows by adding equations (71), the appearing functions of the first and second order term in expansion (70) can be simplified. We get

\[ X(z, v) = 1 - \tau(v)\kappa(v)\sqrt{1 - \frac{z}{\rho(v)}} + \hat{c}(v) \left(1 - \frac{z}{\rho(v)}\right) + O\left(1 - \frac{z}{\rho(v)}\right)^{3/2}. \quad (73) \]

Singularity analysis [5, Proposition 1 and Theorem 1] applied to (73) leads to the uniform expansion

\[ [z^n]X(z, v) = \frac{\tau(v)\kappa(v)}{2\sqrt{\pi}} n^{-\frac{3}{2}} \rho(v)^{-n} (1 + O(n^{-1})), \quad (74) \]

which yields via (17) the following expansion of the moment generating function:

\[ E(e^{P_n s}) = \exp \left( U(s)n + V(s) \right) (1 + O(n^{-1})), \quad (75) \]

with \( U(s) := \log \left( \frac{\rho(ze^{\frac{s}{2}})}{\rho(e^{\frac{s}{2}})} \right) \) and \( V(s) := \log \left( \frac{\tau(e^{\frac{s}{2}})\kappa(e^{\frac{s}{2}})}{\kappa} \right). \)

Now we can apply Theorem 2 to show the asymptotic normality of \( P_n \) as stated in Theorem 3. For the leading term of the expectation resp. the variance we have to compute

\[ U'(0) = \frac{\rho'(1)}{2\rho(1)}, \quad \text{and} \quad U''(0) = \frac{\rho''(1) - \rho'(1)^2 - \rho''(1)}{4\rho^2}, \quad (76) \]

where \( \rho'(1) \) and \( \rho''(1) \) can be obtained from the system of equations (71) by differentiating w. r. t. \( v \) and evaluating at \( v = 1 \). We carry out this only for the quantity \( \rho'(1) \) and omit the lengthy computations for \( \rho''(1) \).
To obtain $\rho'(1)$ we compute $\tau'(1)$ at first: subtracting the equations (71) leads to the equation

$$\rho(v)(v - 1)e^{-\tau(v)}\exp\left(\sum_{k \geq 2} \frac{Y(\rho(v)^k, v^k)}{k}(-1)^k\right) + 1 - \tau(v) = 0, \tag{77}$$

and differentiating (77) gives

$$\rho e^{-\tau} \exp\left(\sum_{k \geq 2} \frac{T(\rho^k)}{k}(-1)^k\right) - \tau'(1) = 0,$$

resp.

$$\tau'(1) = \rho \exp\left(\sum_{k \geq 1} \frac{T(\rho^k)}{k}(-1)^k\right) = \frac{T(\rho^2)}{T(\rho)} = T(\rho^2), \tag{78}$$

where we used $T(\rho) = \tau = 1$ and the identity

$$z \exp\left(\sum_{k \geq 1} \frac{T(z^k)}{k}(-1)^k\right) = \frac{T(z^2)}{T(z)},$$

which follows by applying (6) twice (see [8]).

Differentiating (72) and evaluating at $v = 1$ gives the equation

$$\rho'(1) = -\rho(T(\rho^2) + 1 + 2 \sum_{k \geq 2} E(\rho^k)) \quad \frac{1}{2(1 + \sum_{k \geq 2} T'(\rho^k)\rho^k)}, \tag{79}$$

where $E(z)$ is defined via

$$E(z) := \frac{\partial}{\partial v} Y(z, v) \bigg|_{v=1}.$$

We will not obtain an explicit equation for $E(z)$, but differentiating equation (64a) w. r. t. $v$ and evaluating at $v = 1$ gives the following functional equation for $E(z)$:

$$E(z) = T(z) \sum_{k \geq 1} E(z^k) + \frac{1}{2} T(z) + \frac{1}{2} \frac{T(z^2)}{T(z)}. \tag{80}$$

Therefore equations (76) and (79) lead to the formula

$$U'(0) = \frac{T(\rho^2) + 1 + 2 \sum_{k \geq 2} E(\rho^k)}{4(1 + \sum_{k \geq 2} T'(\rho^k)\rho^k)}, \tag{81}$$

which was already obtained in [8] and evaluated numerically as $U'(0) \approx 0.313456$.

9. Unrooted unlabelled trees

Here we denote by $\tilde{P}_n$ the random variable that counts the path edge-covering number of a randomly chosen unrooted unlabelled tree of size $n$, whereas $P_n$ should denote the corresponding parameter for rooted unlabelled trees. In order to study $\tilde{P}_n$, Meir and Moon [8] used Otter’s dissimilarity characteristic theorem. Let $T$ an unrooted tree of size $n$ constructed by joining the roots of two rooted unlabelled trees $A$ and $B$. Then the number of unrooted trees of size $n$ equals the number of rooted unlabelled trees of size $n$ minus the number of trees $T$ described above in which $A \neq B$. Of course, this immediately leads to
the equation (8), where the factor $\frac{1}{2}$ appearing there is present, because the order of the subtrees $A$ and $B$ is immaterial.

But from this characterization we also obtain immediately the following functional equation for the bivariate generating function $\tilde{X}(z, v) := \sum_{n \geq 0} \sum_{m \geq 0} \hat{T}_n [X_n = m] z^n v^m$, where $\hat{T}_n$ counts the number of nodes of odd degree in random size $n$ unrooted unlabelled trees:

$$\tilde{X}(z, v) = X(z, v) - \frac{1}{2} (Y(z, v)^2 - Y(z^2, v^2)).$$

(82)

where $X(z, v)$ resp. $Y(z, v)$ are the generating functions for rooted unlabelled trees as defined in Section 8. In [8] a different, but equivalent equation was given for $\tilde{X}(z, v)$.

Since $Y(z^2, v^2)$ is analytic for $z$ close to $\rho$ and $v$ close to 1 (see Section 8), we obtain from (82) and the representations (66) and (69) that $\tilde{X}(z, v)$ has in a neighbourhood of $v = 1$ its unique dominant singularity at $z = \rho(v)$, where $\rho(v)$ satisfies the system of functional equations (71). We immediately get the representation

$$\tilde{X}(z, v) = g(z, v) - h(z, v) \sqrt{1 - \frac{z}{\rho(v)}},$$

(83)

with functions $g(z, v)$ and $h(z, v)$ analytic around $z = \rho(v)$ and $v = 1$. But due to the expansions (67) and (73), it follows that the second order term in the expansion of (83) vanishes and we get (uniformly around $v = 1$):

$$\tilde{X}(z, v) = 1 - \frac{1}{2} \tau(v)^2 + \frac{1}{2} Y(\rho(v)^2, v^2) + c(v) \left(1 - \frac{z}{\rho(v)}\right)$$

$$- \tilde{\kappa}(v) \left(1 - \frac{z}{\rho(v)}\right)^2 + \tilde{d}(v) \left(1 - \frac{z}{\rho(v)}\right)^2 + O\left(1 - \frac{z}{\rho(v)}\right)^2),$$

(84)

with certain functions $c(v)$, $\tilde{d}(v)$ and $\tilde{\kappa}(v)$.

Singularity analysis [5, Proposition 1 and Theorem 1] applied to (84) leads to the asymptotic expansion

$$[z^m] \tilde{X}(z, v) = \frac{3\tilde{\kappa}(v)}{4\sqrt{\pi}} n^{-\frac{3}{2}} \rho(v)^{-n} (1 + O(n^{-1})),$$

(85)

which yields via (17) the following expansion of the moment generating function:

$$\mathbb{E}(e^{P_n s}) = \exp \left( U(s) n + V(s) \right) \left(1 + O(n^{-1})\right),$$

(86)

with $U(s) := \log \left( \frac{\rho(s^2)}{\rho(e^s)} \right)$ and $V(s) := \log \left( \frac{\tilde{\kappa}(e^s)}{\tilde{\kappa}} \right)$, where $\tilde{\kappa} := \tilde{\kappa}(1)$.

Therefore we also obtain asymptotic normality of $\hat{P}_n$ and since the functions $U(s)$ appearing in (75) and (86) are equal we obtain that the first order terms of the expectation and the variance of rooted and unrooted unlabelled trees coincide: $\mathbb{E}(\hat{P}_n) \sim \mathbb{E}(P_n)$ and $\mathbb{V}(\hat{P}_n) \sim \mathbb{V}(P_n)$.

Acknowledgment. The author thanks the anonymous referees for many valuable comments.
References


Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Wiedner Hauptstraße 8–10/104, A-1040 Wien, Austria.
E-mail address: Alois.Panholzer@tuwien.ac.at