Combinatorics, Probability and Computing (2008) 17, 21–59. © 2007 Cambridge University Press
doi:10.1017/S0963548307008425 First published online 28 September 2007 Printed in the United Kingdom

The Distribution of Patterns in Random Trees

FRÉDÉRIC CHYZAK,[1] MICHAEL DRMOTA,[2]
THOMAS KLAUSNER[2] and GERARD KOK[3†]

[1]INRIA-Rocquencourt, F-78153 Le Chesnay cedex, France
(e-mail: frederic.chyzak@inria.fr)

[2]Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien,
Wiedner Hauptstraße 8-10/113, A-1040 Wien, Austria
(e-mail: michael.drmota@tuwien.ac.at, klausner@dmg.tuwien.ac.at)

[3]Delft Institute of Applied Mathematics, Delft University of Technology,
Mekelweg 4, NL-2628 CD Delft, The Netherlands
(e-mail: gkok@fsmat.at)

Received 4 May 2006; revised 27 November 2006

Let $\mathcal{T}_n$ denote the set of unrooted labelled trees of size n and let $\mathcal{M}$ be a particular (finite, unlabelled) tree. Assuming that every tree of $\mathcal{T}_n$ is equally likely, it is shown that the limiting distribution as n goes to infinity of the number of occurrences of $\mathcal{M}$ is asymptotically normal with mean value and variance asymptotically equivalent to μn and $\sigma^2 n$, respectively, where the constants $\mu > 0$ and $\sigma \geqslant 0$ are computable.

1. Introduction

In this paper we consider unrooted labelled trees and analyse the number of occurrences of a given tree pattern. More precisely, let $\mathcal{T}_n$ denote the probability space of all unrooted labelled trees of size n with uniform distribution, that is, every tree in $\mathcal{T}_n$ is equally likely. It is, for example, well known that a typical tree in $\mathcal{T}_n$ has about $\mu_k n$ nodes of degree k, where $\mu_k = 1/e(k-1)!$. Moreover, for any fixed k the average number of nodes of degree k in trees in $\mathcal{T}_n$ satisfies a central limit theorem with mean and variance asymptotically equivalent to $\mu_k n$ and $\sigma_k^2 n$ (for a specific constant $\sigma_k > 0$). See [5], where Drmota and Gittenberger explored this phenomenon for unrooted labelled trees and other types of trees.

† This research was supported by the Austrian Science Foundation FWF, grants S8302 and S9604, and by the European AMADEUS project.

A node of degree k is an occurrence of what can be called a star with k edges. In this paper we continue this idea. We consider a pattern $\mathcal{M}$, a given finite tree, and compute the limiting distribution of the number of occurrences of $\mathcal{M}$ in a random member of $\mathcal{T}_n$ as $n \to \infty$. In this context we say that a pattern $\mathcal{M}$ 'occurs' in a tree T, if $\mathcal{M}$ is a subtree of T with the additional property that the node degrees of all internal nodes of $\mathcal{M}$ coincide with the corresponding node degrees of T. Note also that there can be overlaps of two or more copies of $\mathcal{M}$, which we intend to count as separate occurrences.

Our main result in this paper is as follows.

Theorem 1.1. *Let $\mathcal{M}$ be a given finite tree. Then the limiting distribution of the number of occurrences of $\mathcal{M}$ in a random tree of $\mathcal{T}_n$ is asymptotically normal with mean and variance asymptotically equivalent to μn and $\sigma^2 n$, respectively, where $\mu > 0$ and $\sigma^2 \geqslant 0$ depend on the pattern $\mathcal{M}$ and can be computed explicitly and algorithmically and can be represented as polynomials* (*with rational coefficients*) *in $1/e$.*

We consider here a random variable X as Gaussian if its characteristic function is given by $\mathbf{E}\, e^{itX} = e^{i\mu t - \sigma^2 t^2/2}$, that is, the case of zero variance $\sigma^2 = 0$ is included here. For example, if $\mathcal{M}$ consists just of one edge (and two nodes), then the number of occurrences of $\mathcal{M}$ in $\mathcal{T}_n$ is $n-1$ and thus constant. So in that particular case we have $\mu = 1$ and $\sigma^2 = 0$. Nevertheless we conjecture that $\sigma^2 > 0$ in all other cases.

As already mentioned, the case of stars (or nodes of given degree) has been discussed in [5] for various classes of trees. Some previous work for unlabelled trees is due to Robinson and Schwenk [16]. Patterns in (rooted planar) trees have also been considered by Dershowitz and Zaks [6] under the limitation that patterns start at the root. In a work on patterns in random binary search trees, Flajolet, Gourdon and Martínez [7] obtained a central limit theorem. Flajolet and Steyaert also analysed an algorithm for pattern matchings in trees [9, 10, 18]. Further, Ruciński [17] established conditions for when the number of occurrences of a given subgraph in random graphs follows a normal distribution.

The plan of the paper is as follows. In Section 2 we give a short introduction to counting trees with generating functions, and also expand this to two variables for counting stars (nodes of specific degree k) in trees. In Section 3 we expand this framework to the counting of patterns in trees. The resulting asymptotics are presented in Section 4, concluding the proof of Theorem 1.1. Technical details for this as well as explicit algorithms can be found in the appendix. In fact, the algorithmic aspect is one of the driving forces of this paper.

2. Counting trees and counting stars in trees

In this section we introduce a three-step program to count the number of trees in $\mathcal{T}_n$ and in the same fashion the number of occurrences of nodes of degree k of a random tree in $\mathcal{T}_n$. While redundant and probably heavy in this simplistic situation, this procedure was crucial to the derivation in [5] for counting stars and will generalize well to our setting of general tree patterns. We also mention the forthcoming book [8] by Flajolet and Sedgewick as a general reference for combinatorical constructions of that kind and for the singularity analysis of the corresponding generating functions.

We make use, too, of the sets $\mathscr{R}_n$ of rooted labelled trees of size n and $\mathscr{P}_n$ of planted labelled trees of size n. For rooted and unrooted trees, the size n counts the total number of nodes, whether internal or at the leaves. On the other hand, a planted tree is just a rooted tree where the root is adjoined an additional 'phantom' node which does not contribute to the size of the tree, whereas the degree of the root is increased by one. Also, one can think of a planted tree as a rooted tree with an additional edge having no end vertex. The advantage of using planted trees, although it seems to add complexity, will be explained below. Obviously $|\mathscr{P}_n| = |\mathscr{R}_n|$ and $|\mathscr{T}_n| = |\mathscr{R}_n|/n$. It is also well known that $|\mathscr{R}_n| = n^{n-1}$ and $|\mathscr{T}_n| = n^{n-2}$ (see [8]).

The three-step program is the following one. First, the generating function enumerating planted trees is determined, then it is used to count rooted trees by deriving their generating function, and finally the generating function counting unrooted trees is computed.

We define

$$p(x) = \sum_{n=0}^{\infty} |\mathscr{P}_n| \frac{x^n}{n!}, \qquad r(x) = \sum_{n=0}^{\infty} |\mathscr{R}_n| \frac{x^n}{n!}, \qquad t(x) = \sum_{n=0}^{\infty} |\mathscr{T}_n| \frac{x^n}{n!}$$

and proceed in the following way.

(1) **Planted rooted trees.** A planted tree is a planted root node with zero, one, two,... planted subtrees of any order. In terms of the generating function this yields

$$p(x) = \sum_{n=0}^{\infty} \frac{xp(x)^n}{n!} = xe^{p(x)}.$$

(2) **Rooted trees.** For rooted trees we get the same (except for the phantom nodes which are not present here), just a root with zero, one, two,... planted subtrees of any order

$$r(x) = \sum_{n=0}^{\infty} \frac{xp(x)^n}{n!} = xe^{p(x)} = p(x).$$

(3) **Unrooted trees.** Finally, we have $|\mathscr{T}_n| = |\mathscr{R}_n|/n$, as already mentioned. However, we can also express $t(x)$ by a relation which follows from a natural bijection between rooted trees on the one hand and unrooted trees and pairs of planted rooted trees (which are joined by identifying the additional edges at their planted roots and discarding the phantom nodes) on the other hand.[1] This yields

$$t(x) = r(x) - \frac{1}{2}p(x)^2.$$

The functional equation for $p(x)$ can be used either to extract the explicit number $|\mathscr{P}_n| = n^{n-1}$ via Lagrange inversion or to obtain the radius of convergence and asymptotic

[1] Consider the class of rooted (labelled) trees. If the root is labelled by 1 then consider the tree as an unrooted tree. If the root is not labelled by 1 then consider the first edge of the path between the root and 1 and cut the tree into two planted rooted trees at this edge.

expansions of the singular behaviour of this function. It is well known that $x_0 = 1/e$ is the common radius of convergence of $p(x)$, $r(x)$, and $t(x)$, and that the singularity at $x = x_0$ is of square root type:

$$p(x) = r(x) = 1 - \sqrt{2}\sqrt{1-ex} + \frac{2}{3}(1-ex) + \cdots,$$
$$t(x) = \frac{1}{2} - (1-ex) + \frac{2\sqrt{2}}{3}(1-ex)^{3/2} + \cdots.$$

This is reflected by the asymptotic expansions of the numbers

$$|\mathscr{P}_n| = |\mathscr{R}_n| = n^{n-1} \sim \frac{n!}{\sqrt{2\pi}} e^n n^{-3/2},$$
$$|\mathscr{T}_n| = n^{n-2} \sim \frac{n!}{\sqrt{2\pi}} e^n n^{-5/2}.$$

In order to demonstrate the usefulness of the three-step procedure above we repeat the same steps for counting stars with k edges in trees, that is, the number of nodes of degree k, a given fixed positive number. Let $p_{n,m}$ denote the number of planted trees of size n with exactly m nodes of degree k. Furthermore, let $r_{n,m}$ and $t_{n,m}$ be the corresponding numbers for rooted and unrooted trees and set

$$p(x,u) = \sum_{n,m=0}^{\infty} p_{n,m} \frac{x^n u^m}{n!}, \qquad r(x,u) = \sum_{n,m=0}^{\infty} r_{n,m} \frac{x^n u^m}{n!}, \qquad t(x,u) = \sum_{n,m=0}^{\infty} t_{n,m} \frac{x^n u^m}{n!}.$$

Then we have the following (compare with [5]).

(1) **Planted rooted trees.**

$$p(x,u) = \sum_{\substack{n=0 \\ n \neq k-1}}^{\infty} \frac{xp(x,u)^n}{n!} + \frac{xup(x,u)^{k-1}}{(k-1)!} = xe^{p(x,u)} + \frac{x(u-1)p(x,u)^{k-1}}{(k-1)!}.$$

(2) **Rooted trees.**

$$r(x,u) = \sum_{\substack{n=0 \\ n \neq k}}^{\infty} \frac{xp(x,u)^n}{n!} + \frac{xup(x,u)^k}{k!} = xe^{p(x,u)} + \frac{x(u-1)p(x,u)^k}{k!}.$$

(3) **Unrooted trees.** In a similar way to the above we have $t_{n,m} = r_{n,m}/n$, which is sufficient for our purposes. However, as above, it is also possible to express $t(x,u)$ by

$$t(x,u) = r(x,u) - \frac{1}{2}p(x,u)^2.$$

Note that the use of the notion of planted trees is crucial in order to keep track of the nodes of degree k by means of the recursive structure of planted trees. In [5] this approach was used to show that the asymptotic distribution of the number of nodes of degree k in trees of size n is normal, with expectation and variance proportional to n.

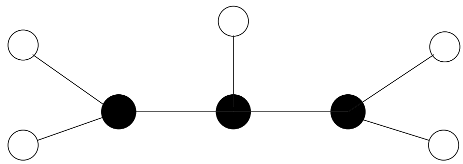

Figure 1. Example pattern.

3. Counting patterns in trees

We now generalize the counting procedure of Section 2 to more complicated patterns. For our purpose, a pattern is a given (finite unrooted unlabelled) tree $\mathcal{M}$. To ease explanations, we will use as $\mathcal{M}$ the example graph in Figure 1.

Recall that we say that a specific pattern $\mathcal{M}$ occurs in a tree T if $\mathcal{M}$ occurs in T as a subtree in the sense that the node degrees for the internal (filled) nodes in the pattern match the degrees of the corresponding nodes in T, while the external (empty) nodes match nodes of arbitrary degree.[2] Because the results for the patterns consisting of only one node or two nodes and one edge are trivial, we now concentrate on patterns with at least three nodes.

Our first aim is to get relations for the generating functions which count the number of occurrences of a specific pattern $\mathcal{M}$. Let $p_{n,m}$ denote the number of planted rooted trees with n nodes and exactly m occurrences of the pattern $\mathcal{M}$ and let

$$p = p(x,u) = \sum_{n,m=0}^{\infty} p_{n,m} \frac{x^n u^m}{n!}$$

be the corresponding generating function.

3.1. Generating functions for planted rooted trees

Proposition 3.1. (Planted rooted trees) *Let $\mathcal{M}$ be a pattern. Then there exists a certain number $L+1$ of auxiliary functions $a_j(x,u)$ $(0 \leqslant j \leqslant L)$ with*

$$p(x,u) = \sum_{j=0}^{L} a_j(x,u)$$

and polynomials $P_j(y_0,\ldots,y_L,u)$ $(1 \leqslant j \leqslant L)$ with non-negative coefficients such that

$$\begin{aligned} a_0(x,u) &= xe^{a_0(x,u)+\cdots+a_L(x,u)} - x\sum_{j=1}^{L} P_j(a_0(x,u),\ldots,a_L(x,u),1) \\ a_1(x,u) &= x \cdot P_1(a_0(x,u),\ldots,a_L(x,u),u) \\ &\vdots \\ a_L(x,u) &= x \cdot P_L(a_0(x,u),\ldots,a_L(x,u),u). \end{aligned} \tag{3.1}$$

[2] More generally, we could also consider pattern-matching problems for patterns in which some degrees of certain possibly external 'filled' nodes must match exactly while the degrees of the other, possibly internal 'empty' nodes might be different. But then the situation is more involved: see Section 5.

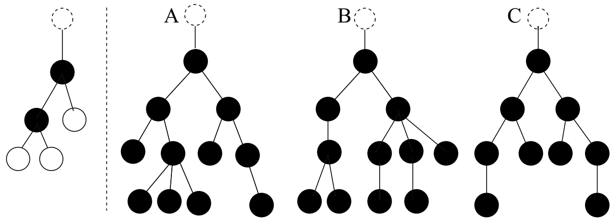

Figure 2. Planted pattern matching.

Furthermore,

$$\sum_{j=1}^{L} P_j(y_0,\dots,y_L,1) \leqslant_c e^{y_0+\cdots+y_L},$$

where $f \leqslant_c g$ means that all Taylor coefficients of the left-hand side are smaller than or equal to the corresponding coefficients of the right-hand side. Moreover, the dependency graph of this system is strongly connected.[3]

The proof of this proposition is in fact the core of the paper. In order to make the arguments more transparent we will demonstrate them with the help of the example pattern in Figure 1. At each step of the proof we will also indicate how to make all constructions explicit so that it is possible to generate system (3.1) effectively.

In a first step we introduce the notion of a *planted pattern*. A planted pattern $\mathscr{M}_p$ is just a planted rooted tree where we again distinguish between internal (filled) and external (empty) nodes. It matches a planted rooted tree from $\mathscr{T}_n$ if $\mathscr{M}_p$ occurs as a proper subtree starting from the (planted) root, that is, the branch structure and node degrees of the filled nodes match. Two occurrences may overlap. For example, in Figure 2 the planted pattern $\mathscr{M}_p$ on the left matches the planted tree A twice (following the left, resp. the right edge from the root), but B not at all. Also remark that, notwithstanding the symmetry of C, the pattern $\mathscr{M}_p$ really matches C twice, as we are interested in matches in labelled trees.

We now construct a planted pattern for each internal (filled) node of our pattern $\mathscr{M}$ which is adjacent to an external (empty) node. The internal (filled) node is considered as the planted root and one of the free attached leaves as the plant. In our example we obtain the two graphs in Figure 3.

The next step is to partition all planted trees according to their degree distribution up to some adequate level. To this end, let D denote the set of out-degrees that occur in the planted patterns introduced above and let h be the maximal height of these patterns. In our example we have $D = \{2\}$ and $h = 3$. For obtaining a partition, we more precisely consider all trees of height less than or equal to h with out-degrees in D. We distinguish two types of leaves in these trees, depending on the depth at which they appear: leaves in level h, denoted '$\circ$', and leaves at levels less than h, denoted '$\Box$'. For our example we get 11 different trees $a_0, a_1, \dots, a_{10}$, depicted in Figure 4.

[3] The notion of dependency graph is explained in Appendix B and, intuitively speaking, reflects the fact that no subsystem can be solved before the whole system.

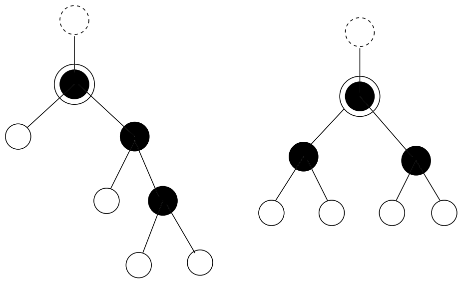

Figure 3. Planted patterns for the pattern in Figure 1.

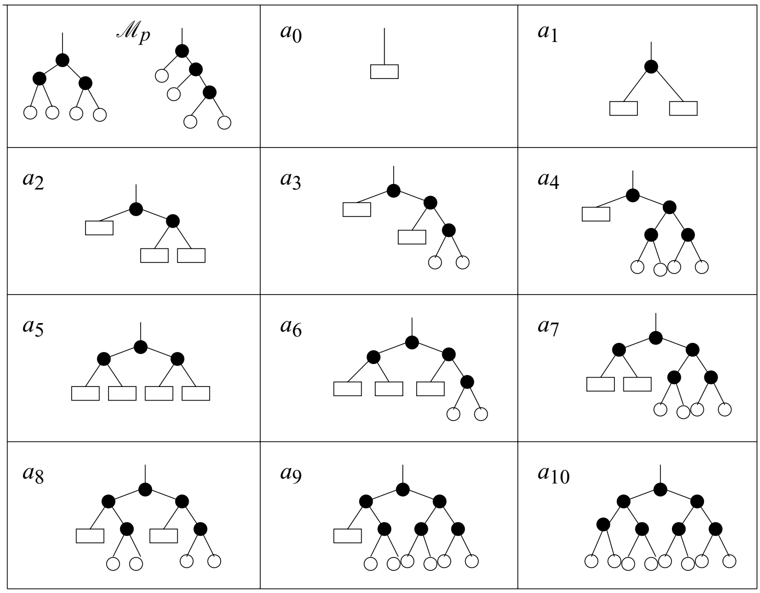

Figure 4. Tree partition.

These trees induce a natural partition of all planted trees for the following interpretation of the two types of leaves. We say that a tree T is contained in class[4] a_j if it matches the finite tree (or pattern) a_j in such a way that a node of type $\square$ has degree not in D, while a node of type $\circ$ has any degree. For example, a_0 corresponds to those planted trees where the out-degree of the root is not in D.

It is easy to observe that these (obviously disjoint) classes of trees form a partition. Indeed, take any rooted tree. For any path from the root to a leaf, consider the first node with out-degree not in D, and replace the whole subtree at it with $\square$. Then replace any node at depth h with $\circ$. The tree obtained in this way is one in the list.

Furthermore, the classes above can be described recursively. To this end, it proves convenient to introduce a formal notation to describe operations between classes of trees: $\oplus$ denotes the disjoint union of classes; $\setminus$ denotes set difference; recursive descriptions of

[4] By abuse of notation the tree class corresponding to the finite tree a_j is denoted by the same symbol a_j.

tree classes are given in the form $a_i = xa_{j_1}^{e_1} \cdots a_{j_\ell}^{e_\ell}$, to express that the class a_i is constructed by attaching e_1 subtrees from the class a_{j_1}, e_2 subtrees from the class a_{j_2}, *etc.*, to a root node that we denote x.

In our example we get the following relations:

$$
\begin{aligned}
a_0 &= p \setminus \bigoplus_{i=1}^{10} a_i = x \oplus x \bigoplus_{i=0}^{10} a_i \oplus x \bigoplus_{n=3}^{\infty} \left(\bigoplus_{i=0}^{10} a_i \right)^n, \\
a_1 &= xa_0^2, \\
a_2 &= xa_0a_1, \\
a_3 &= xa_0(a_2 \oplus a_3 \oplus a_4), \\
a_4 &= xa_0(a_5 \oplus a_6 \oplus a_7 \oplus a_8 \oplus a_9 \oplus a_{10}), \\
a_5 &= xa_1^2, \\
a_6 &= xa_1(a_2 \oplus a_3 \oplus a_4), \\
a_7 &= xa_1(a_5 \oplus a_6 \oplus a_7 \oplus a_8 \oplus a_9 \oplus a_{10}), \\
a_8 &= x(a_2 \oplus a_3 \oplus a_4)^2, \\
a_9 &= x(a_2 \oplus a_3 \oplus a_4)(a_5 \oplus a_6 \oplus a_7 \oplus a_8 \oplus a_9 \oplus a_{10}), \\
a_{10} &= x(a_5 \oplus a_6 \oplus a_7 \oplus a_8 \oplus a_9 \oplus a_{10})^2.
\end{aligned}
$$

This is to be interpreted as follows. Trees in a_1 consist of a (planted) root that is denoted by x that has out-degree 2, and two children that are of out-degree distinct from 2, that is, in a_0. Similarly, trees in a_3 consist of a root x with out-degree 2 and subject to the following additional constraints: one subtree at the root is exactly of type a_0; the other subtree, call it T, is of out-degree 2, either with both subtrees of degree other than 2 (leading to T in a_2), or with one subtree of degree 2 and the other of degree other than 2 (leading to T in a_3), or with both of its subtrees of degree 2 (leading to T in class a_4). Summarizing: $a_3 = xa_0(a_2 \oplus a_3 \oplus a_4)$. Of course this can also be interpreted as $a_3 = xa_0a_2 \oplus xa_0a_3 \oplus xa_0a_4$. Another more involved example corresponds to a_8; here both subtrees are of the form $a_2 \oplus a_3 \oplus a_4$.

To show that the recursive description can be obtained easily in general, consider a tree a_j obtained from some planted pattern $\mathscr{M}_p$. Let $s_1, \ldots, s_d$ denote its subtrees at the root. Then, in each s_i, leaves of type $\circ$ can appear only at level $h-1$. Substitute for all such $\circ$ either $\square$ or a node of out-degree chosen from D and having $\circ$ for all its subtrees. Do this substitution in all possible ways. The collection of trees obtained are some of the a_ks, say $a_{k_1^{(j)}}$, $a_{k_2^{(j)}}$, *etc.* Thus, we obtain the recursive relation

$$a_j = x\big(a_{k_1^{(1)}} \oplus a_{k_2^{(1)}} \oplus \cdots\big) \cdots \big(a_{k_1^{(d)}} \oplus a_{k_2^{(d)}} \oplus \cdots\big)$$

for a_j.

In general, we obtain a partition of $L+1$ classes $a_0, \ldots, a_L$ and corresponding recursive descriptions, where each tree type a_j can be expressed as a disjoint union of tree classes of the kind

$$xa_{j_1} \cdots a_{j_r} = xa_0^{l_0} \cdots a_L^{l_L}, \tag{3.2}$$

where r denotes the degree of the root of a_j and the non-negative integer l_i is the number of repetitions of the tree type a_i.

We proceed to show that this directly leads to a system of equations of the form (3.1), where each polynomial relation stems from a recursive equation between combinatorial classes.

Let Λ_j be the set of tuples $(l_0, \ldots, l_L)$ with the property that $(l_0, \ldots, l_L) \in \Lambda_j$ if and only if the term of type (3.2) is involved in the recursive description of a_j (in expanded form). Further, let $k = K(l_0, \ldots, l_L)$ denote the number of *additional occurrences* of the pattern $\mathcal{M}$ in (3.2) in the following sense: if $b = xa_{j_1} \cdots a_{j_r}$ and T is a (planted rooted) labelled tree of b with subtrees $T_1 \in a_{j_1}$, $T_2 \in a_{j_2}$, *etc.*, and $\mathcal{M}$ occurs m_1 times in T_1, m_2 times in T_2, *etc.*, then T contains $\mathcal{M}$ exactly $m_1 + m_2 + \cdots + m_r + k$ times. The number k corresponds to the number of occurrences of $\mathcal{M}$ in T in which the root of T occurs as internal node of the pattern. By construction of the classes a_i this number only depends on b and not on the particular tree $T \in b$. Let us clarify the calculation of $k = K(l_0, \ldots, l_L)$ with an example. Consider the class a_9 of the partition for the example pattern. Now, in order to determine the number of additional occurrences, we match the planted patterns of Figure 3 at the root of an arbitrary tree of class a_9. The left planted pattern of Figure 3 matches three times, the right one matches once. Thus we find that in this case $k = 4$. For the other classes we find the following values of $k = K(l_0, \ldots, l_L)$:

terms of class	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
value of k	0	0	0	1	2	1	2	3	3	4	5

Now define series P_j by

$$P_j(y_0, \ldots, y_L, u) = \sum_{(l_0, \ldots, l_L) \in \Lambda_j} \frac{1}{l_0! \cdots l_L!} y_0^{l_0} \cdots y_L^{l_L} u^{K(l_0, \ldots, l_L)}.$$

These are in fact polynomials for $1 \leqslant j \leqslant L$ by the finiteness of the corresponding Λ_j. All matches of the planted patterns are handled in the P_j, $1 \leqslant j \leqslant L$, thus

$$P_0(y_0, \ldots, y_L, u) = e^{y_0 + \cdots + y_L} - \sum_{j=1}^{L} P_j(y_0, \ldots, y_L, 1)$$

does not depend on u.

In our pattern we get, for example, for $P_8(y_0, \ldots, y_{10}, u)$

$$\begin{aligned} P_8(y_0, \ldots, y_{10}, u) &= \frac{1}{2} x y_2^2 u^3 + x y_2 y_3 u^3 + x y_2 y_4 u^3 + \frac{1}{2} x y_3^2 u^3 + x y_3 y_4 u^3 + \frac{1}{2} x y_4^2 u^3 \\ &= \frac{1}{2} x (y_2 + y_3 + y_4)^2 u^3. \end{aligned}$$

Finally, let $a_{j;n,m}$ denote the number of planted rooted trees of type a_j with n nodes and m occurrences of the pattern $\mathcal{M}$, and set

$$a_j(x, u) = \sum_{n,m=0}^{\infty} a_{j;n,m} \frac{x^n u^m}{n!}.$$

By this definition it is clear that

$$a_j(x,u) = x \cdot P_j\big(a_0(x,u), \dots, a_L(x,u), u\big),$$

because the size of labelled trees is counted by x (exponential generating function) and the occurrences of the patterns is additive and counted by u. Hence, we explicitly obtain the proposed structure of the system of functional equations (3.1).

For the example pattern we arrive at the following system of equations, where we denote the generating function of the class a_i by the same symbol a_i:

$$\begin{aligned}
a_0 &= a_0(x,u) = p - \sum_{i=1}^{10} a_i = x + x\sum_{i=0}^{10} a_i + x\sum_{n=3}^{\infty}\frac{1}{n!}\left(\sum_{i=0}^{10} a_i\right)^n,\\
a_1 &= a_1(x,u) = \frac{1}{2}xa_0^2,\\
a_2 &= a_2(x,u) = xa_0a_1,\\
a_3 &= a_3(x,u) = xa_0(a_2+a_3+a_4)u,\\
a_4 &= a_4(x,u) = xa_0(a_5+a_6+a_7+a_8+a_9+a_{10})u^2,\\
a_5 &= a_5(x,u) = \frac{1}{2}xa_1^2u,\\
a_6 &= a_6(x,u) = xa_1(a_2+a_3+a_4)u^2,\\
a_7 &= a_7(x,u) = xa_1(a_5+a_6+a_7+a_8+a_9+a_{10})u^3,\\
a_8 &= a_8(x,u) = \frac{1}{2}x(a_2+a_3+a_4)^2u^3,\\
a_9 &= a_9(x,u) = x(a_2+a_3+a_4)(a_5+a_6+a_7+a_8+a_9+a_{10})u^4,\\
a_{10} &= a_{10}(x,u) = \frac{1}{2}x(a_5+a_6+a_7+a_8+a_9+a_{10})^2u^5.
\end{aligned}$$

In order to complete the proof of Proposition 3.1 we just have to show that the dependency graph is strongly connected. By construction, $a_0 = a_0(x,u)$ depends on all functions $a_i = a_i(x,u)$. Thus, it is sufficient to prove that every a_i ($1 \leqslant i \leqslant L$) also depends on a_0. For this purpose consider the subtree of $\mathscr{M}$ that was labelled by a_i and consider a path from its root to an empty node. Each edge of this path corresponds to another subtree of $\mathscr{M}$, say $a_{i_2}, a_{i_3}, \dots, a_{i_r}$. Then, by construction of the system of functional equations above, a_i depends on a_{i_2}, a_{i_2} depends on a_{i_3} *etc.* Finally, the root of a_{i_r} is adjacent to an empty node and thus (the corresponding generating function) depends on a_0. This completes the proof of Proposition 3.1.

Note that we obtain a relatively more compact form of this system by introducing

$$\begin{aligned}
b_0 &= b_0(x,u) = a_0(x,u),\\
b_1 &= b_1(x,u) = a_1(x,u),\\
b_2 &= b_2(x,u) = a_2(x,u) + a_3(x,u) + a_4(x,u)\\
b_3 &= b_3(x,u) = a_5(x,u) + a_6(x,u) + a_7(x,u) + a_8(x,u) + a_9(x,u) + a_{10}(x,u),
\end{aligned} \tag{3.3}$$

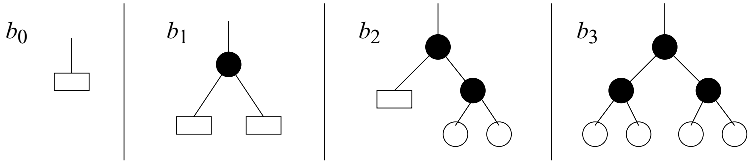

Figure 5. The classes corresponding to the b_i of equations (3.3).

together with the recursive relations

$$b_0 = xe^{b_0+b_1+b_2+b_3} - \frac{1}{2}x(b_0 + b_1 + b_2 + b_3)^2,$$
$$b_1 = \frac{1}{2}xb_0^2,$$
$$b_2 = xb_0b_1 + xb_0b_2u + xb_0b_3u^2,$$
$$b_3 = \frac{1}{2}xb_1^2u + xb_1b_2u^2 + xb_1b_3u^3 + \frac{1}{2}xb_2^2u^3 + xb_2b_3u^4 + \frac{1}{2}xb_3^2u^5.$$

The combinatorial classes corresponding to the b_i (which we will also denote by b_i) have the interpretation shown in Figure 5. We could have obtained the classes b_i directly by restraining the construction to a maximal depth $h - 1$ instead of h. In principle, we could then apply the analytic treatment of Section 4 to the system of the b_i. However, we feel that the existence of a recursive structure of the system of the b_i with a well-defined $K(l_0, .., l_L)$ for each term in the recursive description is slightly less clear. Therefore we preferred to work with the a_i which have a well-defined $K(a_i)$. In Appendix A we will discuss another algorithm that yields in general even more compact systems of equations.

3.2. From planted rooted trees to rooted and unrooted trees

The next step is to find equations for the exponential generating function of rooted trees (where occurrences of the pattern are marked with u). As above we set

$$r(x, u) = \sum_{n,m=0}^{\infty} r_{n,m} \frac{x^n u^m}{n!},$$

where $r_{n,m}$ denotes the number of rooted trees of size n with exactly m occurrences of the pattern $\mathcal{M}$. (That is, occurrences of the *rooted* patterns $\mathcal{M}_r$ deducible from $\mathcal{M}$. Here, a rooted pattern is defined in a very similar way to a planted pattern.)

Proposition 3.2. (Rooted trees) *Let $\mathcal{M}$ be a pattern and let*

$$a_0(x, u), \ldots, a_L(x, u)$$

denote the auxiliary functions introduced in Proposition 3.1. Then there exists a polynomial $Q(y_0, \ldots, y_L, u)$ with non-negative coefficients satisfying $Q(y_0, \ldots, y_L, 1) \leqslant_c e^{y_0+\cdots+y_L}$, and such that

$$r(x, u) = G(x, u, a_0(x, u), \ldots, a_L(x, u)) \tag{3.4}$$

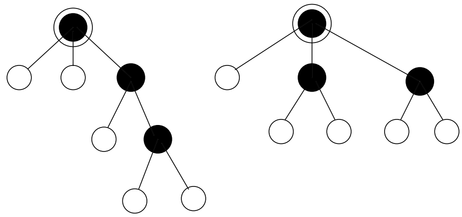

Figure 6. Rooted patterns for the pattern in Figure 1.

for

$$G(x, u, y_0, \dots, y_L) = x\big(e^{y_0+\cdots+y_L} - Q(y_0, \dots, y_L, 1) + Q(y_0, \dots, y_L, u)\big). \qquad (3.5)$$

Proof. The proof is in principle a direct continuation of the proof of Proposition 3.1. We recall that a rooted tree is just a root with zero, one, two, … planted subtrees, *i.e.*, the class of rooted trees can be described as a disjoint union of classes c of rooted trees of the form $xa_{j_1} \cdots a_{j_d}$. Furthermore, let l_i denote the number of classes a_i in this term such that $c = xa_0^{l_0} \cdots a_L^{l_L}$, and set $\bar{K}(l_0, \dots, l_L)$ to be the number of additional occurrences of the pattern $\mathscr{M}$. This number again corresponds to the number of occurrences of $\mathscr{M}$ in a (rooted) tree $T \in c$ in which the root of T occurs as internal node of the pattern. Set

$$Q_d(y_0, \dots, y_L, u) = \sum_{l_0+\cdots+l_L=d} \frac{1}{l_0! \cdots l_L!} y_0^{l_0} \cdots y_L^{l_L} u^{\bar{K}(l_0,\dots,l_L)}.$$

Then, by construction,

$$r(x, u) = x \sum_{d \geqslant 0} Q_d(a_0(x, u), \dots, a_L(x, u), u).$$

Note that $\sum_{d\geqslant 0} Q_d(y_0, \dots, y_L, 1) = e^{y_0+\cdots+y_L}$. Let $\bar{D}$ denote the set of degrees of the internal (filled) nodes of the pattern, that is, $\bar{D} = \{\, d+1 : d \in D \,\}$; then $Q_d(y_0, \dots, y_L, u)$ does not depend on u if $d \notin \bar{D}$. With

$$Q(y_0, \dots, y_L, u) := \sum_{d \in \bar{D}} Q_d(y_0, \dots, y_L, u),$$

we obtain (3.4) and (3.5). The number $\bar{K}(l_0, \dots, l_L)$ is well-defined for a similar reason as was $K(l_0, \dots, l_L)$, and can be calculated similarly. □

We again illustrate the proof with our example. In Figure 6 the corresponding rooted patterns are shown. For convenience let $r_0 = r_0(x, u)$ denote the function

$$r_0 = xe^p - \frac{xp^3}{3!},$$

where $p = a_0 + \cdots + a_{10}$. The function r_0 might also be interpreted as a catch-all function for the ‘uninteresting’ subtrees – just a root x with an unspecified number of planted trees attached, except the ones we handle differently, namely the cases $d \in \bar{D} = \{3\}$. The generating function $r = r(x,u)$ for rooted trees is then given by

$$r = r_0 + \frac{1}{6}xb_0^3 + \frac{1}{2}x \sum_{1\leqslant i\leqslant 3} b_0^2 b_i u^{i-1} + \frac{1}{2}x \sum_{1\leqslant i,j\leqslant 3} b_0 b_i b_j u^{i+j-1} + \frac{1}{6}x \sum_{1\leqslant i,j,k\leqslant 3} b_i b_j b_k u^{i+j+k},$$

where the b_i are defined in (3.3).

As above we have $t_{n,m} = r_{n,m}/n$, where $t_{n,m}$ denotes the number of unrooted trees with n nodes and exactly m occurrences of the pattern $\mathscr{M}$. This relation is sufficient for our purposes. It is also possible to express the corresponding generating function $t(x,u)$. In a way similar as before, we can define the number of additional occurrences $\hat{K}(i,j)$ of the pattern $\mathscr{M}$ that appear by constructing an unrooted tree from two planted trees of the class a_i and a_j by identifying the additional edges at their planted roots and discarding the phantom nodes. For our example we get

$$t(x,u) = r(x,u) - \frac{1}{2}p(x,u)^2 - \frac{1}{2} \sum_{1\leqslant i,j\leqslant 3} b_i(x,u)b_j(x,u)(u^{i+j-2} - 1).$$

4. Asymptotic behaviour

Although the exact number of occurrences of the pattern is encoded in the system of equations (3.1) together with (3.4), this kind of implicit representation cannot be used to obtain (nice) explicit representations for these numbers. However, if we are interested in the asymptotic behaviour, we do not have to compute explicit formulae from the system of equations. Instead, we apply a result slightly adapted from [4] which we state and discuss in Appendix B. In fact, it is immediately clear that Theorem B.1 in this appendix, whose object is the proof of Gaussian limiting distributions, applies to the kind of problem we are interested in: the assertions of Propositions 3.1 and 3.2 exactly fit the assumptions of Theorem B.1.

The only missing point is the existence of a non-negative solution $(x_0, \mathbf{a}_0)$ of the system

$$\mathbf{a} = \mathbf{F}(x, \mathbf{a}, 1), \tag{4.1}$$

$$0 = \det(\mathbf{I} - \mathbf{F_a}(x, \mathbf{a}, 1)), \tag{4.2}$$

where (4.1) is the system of functional equations of Proposition 3.1 and $\mathbf{F_a}$ is the Jacobian matrix of $\mathbf{F}$. Since the sum of all unknown functions $p(x,u)$ is known for $u = 1$,

$$p(x,1) = p(x) = \sum_{n\geqslant 1} n^{n-1}\frac{x^n}{n!} = 1 - \sqrt{2}\sqrt{1-ex} + \cdots,$$

it is not unexpected that $x_0 = 1/e$.

Proposition 4.1. *There exists a unique non-negative solution* $(x_0, \mathbf{a}_0)$ *of system* (4.1)–(4.2), *for which* $x_0 = 1/e$ *and the components of* $\mathbf{a}_0$ *are polynomials* (*with rational coefficients*) *in* $1/e$.

Proof. For a proof, set $u = 1$ and consider the solution $\mathbf{a}(x,1) = (a_0(x,1), \ldots, a_L(x,1))$. Since the dependency graph is strongly connected it follows that all functions $a_j(x,1)$ have the same radius of convergence, which has to be $x_0 = 1/e$, and all functions are singular at $x = x_0$. Since $0 \leqslant a_j(x,1) \leqslant p(x,1) < \infty$ for $0 \leqslant x \leqslant x_0$, it also follows that $a_j(x_0,1)$ is finite, and we have $\mathbf{a}(x_0,1) = \mathbf{F}(x_0, \mathbf{a}(x_0,1),1)$. If we had the inequality $\det(\mathbf{I} - \mathbf{F}_{\mathbf{a}}(x_0, \mathbf{a}(x_0,1),1)) \neq 0$ then the implicit function theorem would imply the existence of an analytic continuation for $a_j(x,1)$ around $x = x_0$, which is, of course, a contradiction. Thus, the determinant is zero and system (4.1)–(4.2) has a unique solution.

To see that the components $\bar{a}_0, \ldots, \bar{a}_L$ (with $\bar{a}_i = a_i(1/e,1)$) of $\mathbf{a}_0$ are polynomials in $1/e$ we will construct the partition $\mathscr{A} = \{a_0, a_1, \ldots, a_L\}$ on which the system of equations (4.1)–(4.2) is based, by refining step by step the trivial partition consisting of only one class p. The recursive description of this trivial partition is given by the formal equation $p = x \sum_{i \geqslant 0} p^i$. Additionally, the solution of the corresponding equation $p = x\exp(p)$ for the generating function p (denoted by the same symbol p) is given by $(x_0, \bar{p}) = (1/e, 1)$, with $\bar{p}$ clearly a (constant) polynomial in $1/e$. Now let $D = \{d_1, \ldots, d_s\}$ $(s \in \mathbb{N})$ again denote the set of out-degrees that occur in the planted patterns. We will refine p by introducing for each $d_i \in D$ a class a_i consisting of all trees of root out-degree d_i, as well as a class a_0 for trees with root out-degree not in D. The partition $\{a_0, a_1, \ldots, a_s\}$ has the recursive description

$$\begin{aligned} a_0 &= x \sum_{j \in \mathbb{N} \setminus D} (a_0 \oplus a_1 \oplus \cdots \oplus a_s)^j, \\ a_i &= x(a_0 \oplus a_1 \oplus \cdots \oplus a_s)^{d_i} \qquad (i = 1, \ldots, s), \end{aligned} \tag{4.3}$$

and the solution of the corresponding system of equations

$$\begin{aligned} a_0(x,1) &= x \sum_{j \in \mathbb{N} \setminus D} \frac{1}{j!} (a_0(x,1) + a_1(x,1) + \cdots + a_s(x,1))^j \\ &= x e^{a_0(x,1) + \cdots + a_s(x,1)} - x \sum_{i=1}^{s} \frac{1}{d_i!} (a_0(x,1) + a_1(x,1) + \cdots + a_s(x,1))^{d_i} \\ &= x e^{p(x)} - x \sum_{i=1}^{s} \frac{1}{d_i!} p(x)^{d_i}, \\ a_i(x,1) &= \frac{x}{d_i!} (a_0(x,1) + a_1(x,1) + \cdots + a_s(x,1))^{d_i} = \frac{x}{d_i!} p(x)^{d_i} \qquad (i = 1, \ldots, s), \end{aligned} \tag{4.4}$$

is given by

$$x_0 = 1/e, \qquad \bar{a}_i = \frac{1}{d_i!\, e} \quad (i = 1, \ldots, s), \qquad \bar{a}_0 = 1 - (\bar{a}_1 + \cdots + \bar{a}_s), \tag{4.5}$$

thus again polynomials in $1/e$. We continue by refining this last partition by introducing classes $c_1, \ldots, c_m$ (for some $m \in \mathbb{N}$) for each term at the right-hand side of (4.3) after expanding the ‘multinomial’. Such a class c_j is of the form

$$c_j = x a_0^{l_0^{(j)}} a_1^{l_1^{(j)}} \cdots a_s^{l_s^{(j)}}$$

with natural numbers $l_i^{(j)}$, $i = 0, \ldots, s$. We get a new partition $\{a_0, c_1, \ldots, c_m\}$ which has a recursive description by construction (because we can replace the a_i by disjoint unions of certain c_j). The corresponding system of equations for the generating functions is given by

$$c_j(x,1) = \frac{x}{l_0^{(j)}!\, l_1^{(j)}! \cdots l_s^{(j)}!} a_0(x,1)^{l_0^{(j)}} a_1(x,1)^{l_1^{(j)}} \cdots a_s(x,1)^{l_s^{(j)}} \quad (j = 1, \ldots, u),$$

and consequently we have for $x_0 = 1/e$ the solution

$$\bar{c}_j = \frac{1}{e}\frac{1}{l_0^{(j)}!\,l_1^{(j)}!\cdots l_s^{(j)}!}\bar{a}_0^{l_0^{(j)}}\bar{a}_1^{l_1^{(j)}}\cdots\bar{a}_s^{l_s^{(j)}} \quad (j = 1,\ldots,m)$$

with the $\bar{a}_i$ of (4.5). Thus the $\bar{c}_j$ are again polynomials in $1/e$. By continuing this procedure until level h (*i.e.*, performing the refinement step h times) we end up with the partition $\mathscr{A}$ and we see that the solution for the corresponding system of equations consists of polynomials in $1/e$, which completes the proof of Proposition 4.1. □

Note that there is a close link with Galton–Watson branching processes. Let $p_k = \frac{1}{k!e}$ denote a Poisson offspring distribution. Now we interpret a class a_i as the class of process realizations for which the (non-planar) branching structure at the beginning of the processes corresponds to the root structure of a_i. Then $\bar{a}_i = a_i(1/e, 1)$ is just the probability of this event.

We now solve the system of equations obtained for the example pattern. We have $x_0 = 1/e$. The components of $\mathbf{a}_0$ can be easily obtained by following the construction of the proof of Proposition 4.2 (or we use the branching process interpretation). For example, if we set $p = 1/(2e)$ for the probability of an out-degree 2 and $q = 1 - p$ then we get $\bar{a}_4 = a_4(1/e, 1) = 2qp^3 = \frac{2e-1}{16e^5}$. The factor 2 comes from the fact that the two subtrees of the root may be interchanged: see Figure 4. The other classes can be treated similarly, and we find:

$$\begin{aligned}
p(1/e,1) &= 1, & a_5(1/e,1) &= (2e-1)^4/(128e^7),\\
a_0(1/e,1) &= (2e-1)/(2e), & a_6(1/e,1) &= (2e-1)^3/(32e^7),\\
a_1(1/e,1) &= (2e-1)^2/(8e^3), & a_7(1/e,1) &= (2e-1)^2/(64e^7),\\
a_2(1/e,1) &= (2e-1)^3/(16e^5), & a_8(1/e,1) &= (2e-1)^2/(32e^7),\\
a_3(1/e,1) &= (2e-1)^2/(8e^5), & a_9(1/e,1) &= (2e-1)/(32e^7),\\
a_4(1/e,1) &= (2e-1)/(16e^5), & a_{10}(1/e,1) &= 1/(128e^7).
\end{aligned} \tag{4.6}$$

We are now ready to complete the proof of the main part of Theorem 1.1. By Propositions 3.1–4.1 we can apply Theorem B.1 and it follows that the numbers $r_{n,m}$ have a Gaussian limiting distribution with mean and variance which are proportional to n. Since $t_{n,m} = r_{n,m}/n$ we get exactly the same law for unrooted trees. It remains to compute μ and σ^2.

By using the procedure described in Appendix B we get for our example pattern

$$\mu = \frac{5}{8e^3} = 0.0311169177\cdots$$

and

$$\sigma^2 = \frac{20e^3 + 72e^2 + 84e - 175}{32e^6} = 0.0764585401\cdots.$$

We observe – as predicted by Theorem 1.1 – that both μ and σ^2 can be written as rational polynomials in $1/e$.

In what follows we will prove this fact (which completes the proof of Theorem 1.1) and also present an easy formula for μ. Unfortunately the procedure for calculating σ^2 is much more complicated so that it seems that there is no simple formula.

Proposition 4.2. *Let $x_0 = 1/e$ and $\mathbf{a}_0$ be given by Proposition 4.1 and let $P_j(\mathbf{y}, u)$ ($1 \leqslant j \leqslant L$) be the polynomials of Proposition 3.1, with $\mathbf{y} = (y_0, \ldots, y_L)$. Then μ (of Theorem 1.1) is a polynomial in $1/e$ with rational coefficients and is given by*

$$\mu = \frac{1}{e} \sum_{j=1}^{L} \frac{\partial P_j}{\partial u}(\mathbf{a}_0, 1). \tag{4.7}$$

Proof. Let $\mathbf{a} = \mathbf{F}(x, \mathbf{a}, u)$ be the system of functional equations of Proposition 3.1. In Appendix B the following formula for the mean is derived:

$$\mu = \frac{1}{x_0} \frac{\mathbf{b}^{\mathrm{T}} \mathbf{F}_u(x_0, \mathbf{a}_0, 1)}{\mathbf{b}^{\mathrm{T}} \mathbf{F}_x(x_0, \mathbf{a}_0, 1)}. \tag{4.8}$$

Here $\mathbf{b}^T$ denotes a positive left eigenvector of $\mathbf{I} - \mathbf{F}_\mathbf{a}$, which is unique up to scaling.

From the equality

$$\mathbf{F}(x, \mathbf{a}, u) = \begin{pmatrix} x\left(e^{a_0 + \cdots + a_L} - \sum_{j=1}^{L} P_j(\mathbf{a}, 1)\right) \\ x P_1(\mathbf{a}, u) \\ x P_2(\mathbf{a}, u) \\ \vdots \\ x P_L(\mathbf{a}, u) \end{pmatrix},$$

we get, after denoting $\frac{\partial P_i}{\partial a_j}$ with P_{i,a_j},

$$\mathbf{F}_\mathbf{a} = x \begin{pmatrix} e^{a_0 + \cdots + a_L} - \sum_{j=1}^{L} P_{j,a_0} & \cdots & e^{a_0 + \cdots + a_L} - \sum_{j=1}^{L} P_{j,a_L} \\ P_{1,a_0} & \cdots & P_{1,a_L} \\ \vdots & & \vdots \\ P_{L,a_0} & \cdots & P_{L,a_L} \end{pmatrix}. \tag{4.9}$$

Since $a_0(x_0, 1) + \cdots + a_L(x_0, 1) = p(x_0, 1) = 1$ we have $x_0 e^{a_0(x_0,1) + \cdots a_L(x_0,1)} = 1$. Consequently the sum of all rows of $\mathbf{F}_\mathbf{a}$ equals $(1, 1, \ldots, 1)$ for $x = x_0 = 1/e$. Thus, denoting the transpose of a vector v by v^{T}, the vector $\mathbf{b}^{\mathrm{T}} = (1, 1, \ldots, 1)$ is the unique positive left eigenvector of $\mathbf{I} - \mathbf{F}_\mathbf{a}$, up to scaling.

It is now easy to check that

$$x_0 \mathbf{b}^{\mathrm{T}} \mathbf{F}_x(x_0, \mathbf{a}_0, 1) = \frac{1}{e} e^{a_0(x_0,1) + \cdots a_L(x_0,1)} = 1$$

and that

$$\mathbf{b}^{\mathrm{T}} \mathbf{F}_u(x_0, \mathbf{a}_0, 1) = \frac{1}{e} \sum_{j=1}^{L} P_{j,u}(\mathbf{a}_0, 1).$$

The fact that μ is a polynomial in $1/e$ is now a direct consequence from the fact that $\mathbf{a}_0$ consists of polynomials in $1/e$ and the fact that the coefficients are rational follows from the fact that $\mathbf{F}(x, \mathbf{a}, u)$ has rational coefficients. □

Of course, with the help of (4.7) we can easily evaluate μ directly. As already indicated it seems that there is no simple formula for σ^2.

Before proving Proposition 4.4 we state an interesting fact that will be used henceforth.

Lemma 4.3. *Let $a_0, a_1, \ldots, a_L$ be the partition of p that is used in the proof of Theorem 1.1. Then*

$$\det\big(\mathbf{I} - \mathbf{F_a}(x, \mathbf{a}, 1)\big) = 1 - xe^{a_0+a_1+\cdots+a_L}.$$

Since the proof is a rather lengthy computation we postpone it to Appendix C.

Proposition 4.4. *Let $x_0 = 1/e$ and $\mathbf{a}_0$ be given by Proposition 4.1. Then σ^2 (of Theorem 1.1) is a polynomial in $1/e$ (with rational coefficients).* □

Proof. From the proof of Proposition 4.2 we already know that $x_u(1)$ can be represented as a polynomial in $1/e$ (with rational coefficients). The next step is to show that $\mathbf{a}_u(1)$ has the same property. For this purpose we have to look at the system (B.11)

$$(\mathbf{I} - \mathbf{F_a})\mathbf{a}_u = \mathbf{F}_x x_u + \mathbf{F}_u,$$
$$-D_\mathbf{a}\mathbf{a}_u = D_x x_u + D_u,$$

where $D(x, \mathbf{a}, u) = \det(\mathbf{I} - \mathbf{F_a}(x, \mathbf{a}, 1)) = 1 - xe^{a_0+a_1+\cdots+a_L}$. We first observe that

$$D_\mathbf{a}(x_0, \mathbf{a}_0, 1) = (-1, -1, \ldots, -1).$$

Hence, we can replace the first row of the $(L+1)\times(L+1)$-matrix $\mathbf{I} - \mathbf{F_a}$ (which is redundant since the matrix has rank L) by the row $(1, 1, \ldots, 1)$ and obtain a regular linear system for $\mathbf{a}_u(1)$. Note that all entries of the right-hand side of this linear system can be represented as polynomials in $1/e$.

Let $\mathbf{M}(x, \mathbf{a})$ denote the matrix obtained from $\mathbf{I} - \mathbf{F_a}(x, \mathbf{a}, 1)$ by replacing the first row by $(1, 1, \ldots, 1)$. It follows from the proof of Lemma 4.3 that $\det \mathbf{M}(x, \mathbf{a}) = 1$. Further, all entries of $\mathbf{M}(x_0, \mathbf{a}_0)$ can be represented as polynomials in $1/e$. Thus, $\mathbf{M}(x_0, \mathbf{a}_0)^{-1}$ has the same property and consequently $\mathbf{a}_u(1)$ has this property too.

From that it directly follows from (B.12) that x_{uu} is also represented as a polynomial in $1/e$. (By definition, $b(x, \mathbf{a}, u)$ is a rational polynomial of the entries of $\mathbf{I} - \mathbf{F_a}$.)

With the help of (B.4) this finally leads to a representation of σ^2 as a polynomial in $1/e$. □

This completes the proof of Theorem 1.1.

5. Extensions and generalizations

In what follows we list some obvious and some less obvious extensions of our main result. For conciseness we do not present the details.

5.1. Several patterns

Let $\mathcal{M}_1, \ldots, \mathcal{M}_k$ be k different patterns. Then the problem is to determine the joint (limiting) distribution of the number of occurrences of $\mathcal{M}_1, \ldots, \mathcal{M}_k$ in trees of size n.

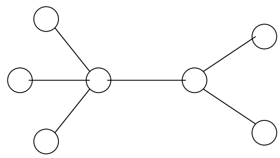

Figure 7. Example pattern with empty nodes.

Using the same techniques as above (introducing the forest of planted patterns deduced from the patterns) we again obtain a system of functional equations. The only difference is that we now have to count occurrences of $\mathscr{M}_1, \ldots, \mathscr{M}_k$ with different variables $u_1, \ldots, u_k$, which is done in the same fashion as for a single u. In view of Theorem B.1, multiple variables u make no difference and we obtain a multivariate Gaussian limiting distribution.

5.2. Patterns containing paths of unspecified length

It might also be interesting to consider patterns where specific edges can be replaced by paths of arbitrary length. It turns out that this case in particular is more involved since a natural partition of all planted rooted trees is now infinite. Nevertheless it is possible to replace infinite series of such classes by one new class and end up with a finite system. Thus, this leads to a Gaussian limit law (as above).

5.3. Filled and empty nodes

In our model we have distinguished between internal (filled) and external (empty) nodes of the pattern $\mathscr{M}$, where the degrees of the internal (filled) nodes have to match exactly. It also seems possible to consider the following more general matching problem. Let $\mathscr{M}$ again be a finite tree, where certain nodes are 'filled' and the remaining ones are 'empty'. Now we say that $\mathscr{M}$ matches if it occurs as a subtree such that the corresponding degrees of the filled nodes are equal whereas the degrees of the empty nodes might be different. It seems that the counting procedure above can be adapted to cover this case, too. However, it is definitely more involved. For example, if leaves of the pattern are filled nodes then these nodes have to be leaves wherever the pattern occurs. This implies that some of the functions $a_j(x, u)$ are then explicitly given in the system and the dependency graph is not strongly connected. However, it seems that this situation can be managed by eliminating these functions. Furthermore, and this is more serious, in general one has to consider infinitely many classes of trees leading to an infinite system of functional equations, in particular if an internal node is 'empty'. In such a case Theorem B.1 cannot be applied any longer. Nevertheless we hope that the approach of Lalley [13], which is applicable to infinite systems of functional equations in one variable, can be generalized to a corresponding generalization of Theorem B.1 to proper infinite systems. Thus, we can expect a Gaussian limit law even in this case.

In order to be more precise we will present an easy example. Let $\mathscr{M}$ denote the pattern depicted in Figure 7. Here all nodes are empty. Thus, the corresponding pattern counting problem is a subgraph counting problem.

We partition all planted trees according to their root degree. Let a_k denote the set of planted rooted trees with root out-degree k and $a_k(x, u)$ the corresponding generating function (which also counts the number of subgraph occurrences of $\mathcal{M}$). Further, let $r(x,u)$ denote the generating function of rooted trees. Then we have

$$a_k(x,u) = \frac{x}{k!}\left(\sum_{i\geqslant 0} a_i(x,u)u^{\binom{k}{2}\binom{i}{3}+\binom{k}{3}\binom{i}{2}}\right)^k \quad (k \geqslant 0)$$

and

$$r(x,u) = x\sum_{k\geqslant 0}\frac{1}{k!}\left(\sum_{i\geqslant 0} a_i(x,u)u^{\binom{k-1}{2}\binom{i}{3}+\binom{k-1}{3}\binom{i}{2}}\right)^k.$$

This system is easy to solve for $u = 1$. Here we have $a_k(x,1) = xp(x)^k/k!$ and $r(x,1) = p(x)$. By taking derivatives with respect to u and summing over all k we also get (after some algebra)

$$r_u(x,1) = \frac{5}{12}\frac{p(x)^7}{1-p(x)} + \frac{1}{6}\frac{p(x)^8}{1-p(x)} + \frac{p(x)^7}{6}.$$

This implies that the average value of pattern occurrences (in this sense) is of the form $(7/12)n + O(1)$, that is, $\mu = 7/12$. In principle it is also possible to get asymptotics for higher moments but the calculations get more and more involved.

5.4. Simply generated trees

Simply generated trees have been introduced by Meir and Moon [14] and are proper generalizations of several types of rooted trees. Let

$$\varphi(x) = \varphi_0 + \varphi_1 x + \varphi_2 x^2 + \cdots$$

be a power series with non-negative coefficients; in particular we assume that $\varphi_0 > 0$ and $\varphi_j > 0$ for some $j \geqslant 2$. We then define the weight $\omega(T)$ of a finite rooted tree T by

$$\omega(T) = \prod_{j\geqslant 0}\varphi_j^{D_j(T)},$$

where $D_j(T)$ denotes the number of nodes in T with j successors. If we set

$$y_n = \sum_{|T|=n}\omega(T),$$

then the generating function

$$y(x) = \sum_{n\geqslant 1} y_n x^n$$

satisfies the functional equation

$$y(x) = x\varphi(y(x)).$$

In this context, y_n denotes a weighted number of trees of size n. For example, if $\varphi_j = 1$ for all $j \geqslant 0$ (that is, $\varphi(x) = 1/(1-x)$) then all rooted trees have weight $\omega(T) = 1$ and $y_n = p_n$ is the number of planted plane trees. If $\varphi_j = 1/j!$ (that is, $\varphi(x) = e^x$) then we formally get labelled rooted trees, *etc.*

Of course, we can proceed in the same way as above and obtain a system of functional equations that counts occurrences of a specific pattern in simply generated trees, and (under suitable conditions on the growth of φ_j) we finally obtain a Gaussian limiting distribution. This has explicitly been done by Kok in his thesis [11, 12].

5.5. Unlabelled trees

Let $\hat{p}_n$ denote the number of unlabelled planted rooted trees and $\hat{t}_n$ the number of unlabelled unrooted trees. The generating functions are denoted by

$$\hat{p}(x) = \sum_{n\geqslant 1} \hat{p}_n x^n \quad \text{and} \quad \hat{t}(x) = \sum_{n\geqslant 1} \hat{t}_n x^n.$$

The structure of these trees is much more difficult than that of labelled trees. It turns out that one has to apply Pólya's theory of counting and an amazing observation (5.1) by Otter [15]. The generating functions $\hat{p}(x)$ and $\hat{t}(x)$ satisfy the functional equations

$$\hat{p}(x) = x \sum_{k\geqslant 0} Z\left(S_k; \hat{p}(x), \hat{p}(x^2), \ldots, \hat{p}(x^k)\right) = x \exp\left(\hat{p}(x) + \frac{1}{2}\hat{p}(x^2) + \frac{1}{3}\hat{p}(x^3) + \cdots\right)$$

and

$$\hat{t}(x) = \hat{p}(x) - \frac{1}{2}\hat{p}(x)^2 + \frac{1}{2}\hat{p}(x^2), \tag{5.1}$$

where $Z(S_k; x_1, \ldots, x_k)$ denotes the cycle index of the symmetric group S_k. These functions have a common radius of convergence $\rho \approx 0.338219$ and a local expansion of the form

$$\hat{p}(x) = 1 - b(\rho - x)^{1/2} + c(\rho - x) + d(\rho - x)^{3/2} + \mathcal{O}\left((\rho - x)^2\right)$$

and

$$\hat{t}(x) = \frac{1 + \hat{p}(\rho^2)}{2} - \frac{b^2 + 2\rho\hat{p}'(\rho^2)}{2}(\rho - x) + bc(\rho - x)^{3/2} + \mathcal{O}\left((\rho - x)^2\right),$$

where $b \approx 2.6811266$ and $c = b^2/3 \approx 2.3961466$, and $x = \rho$ is the only singularity on the circle of convergence $|x| = \rho$. Thus, they behave like $p(x)$ and $t(x)$. We also get

$$\hat{p}_n = \frac{b\sqrt{\rho}}{2\sqrt{\pi}} n^{-3/2} \rho^{-n} \left(1 + \mathcal{O}(n^{-1})\right)$$

and

$$\hat{t}_n = \frac{b^3 \rho^{3/2}}{4\sqrt{\pi}} n^{-5/2} \rho^{-n} \left(1 + \mathcal{O}(n^{-1})\right).$$

Furthermore, it is possible to count the number of nodes of specific degree with the help of bivariate generating functions (compare with [5]). Thus, using Pólya's theory of counting we can also obtain a system of functional equations for bivariate generating functions that count the number of occurrences of a specific pattern. The major difference to the procedure above is that this system also contains terms of the form $a_j(x^k, u^k)$ for $k \geqslant 2$. Fortunately these terms can be considered as known functions when x varies around the singularity ρ and u varies around 1 (compare again with [5]). Hence, Theorem B.1 applies again and we can proceed as above. This has explicitly been done by Kok in his thesis [11, 12].

5.6. Forests

First, let us consider the case of labelled trees with generating function $t(x, u)$. Then the generating function $f(x, u)$ of unlabelled forests is given by

$$f(x, u) = e^{t(x,u)}.$$

Thus, the singular behaviour of $f(x, u)$ is the same as that of $t(x, u)$ (compare with [5]) and consequently we again obtain a Gaussian limiting distribution for the number of occurrences of a specific pattern in labelled forests.

The case of unlabelled forests is similar. Here we have

$$\hat{f}(x, u) = \exp\left(\hat{t}(x, u) + \frac{1}{2}\hat{t}(x^2, u^2) + \frac{1}{3}\hat{t}(x^3, u^3) + \cdots\right).$$

Of course, we can consider other classes of trees or forests of a given number of trees.

5.7. Forbidden patterns

It is also interesting to count the number $t_{n,0}$ of trees of size n without a given pattern. The generating function of these numbers is just $p(x, 0)$, resp. $t(x, 0)$. It is now an easy exercise to show that there exists an $\eta > 0$ such that

$$t_{n,0} \leqslant t_n e^{-\eta n}.$$

The only thing we have to check is that the radius of convergence of $t(x, 0)$ is larger than the radius of convergence of $t(x, 1)$. However, this is obvious since the radius of convergence of $t(x, u)$ (which is the same as that of $p(x, u)$) is given by $x(u)$ (for u around 1) and $x'(1) < 0$.

Appendix A: Algorithms

In the main part of this paper we showed that the limiting distribution of the number of pattern occurrences is normal with computable μ and σ^2. However, the family of classes $\{a_0, a_1, \ldots, a_L\}$ considered in the first part was especially created to make the arguments more transparent; there were no considerations about minimality. In this appendix we focus on creating another partition $\mathscr{A} = \{a_0, \ldots, a_L\}$ of p which has considerably fewer classes. It also has the properties that it is recursively describable and allows an unambiguous definition of the number of additional occurrences $K(l_0, \ldots, l_L)$ of the pattern. For example, we show that for the pattern of Figure 9 we need just 8 equations, whereas the previous proof would use more than 1000 equations.

First we remark that in some cases it is profitable to adjust the structure of the system of equations (3.1) in Proposition 3.1 by allowing an additional polynomial $P_0(y_0, \ldots, y_L, u)$ in the first equation. The first equation then becomes

$$\begin{aligned} a_0(x, u) = {} & x \cdot P_0(a_0(x, u), \ldots, a_L(x, u), u) \\ & + (xe^{a_0(x,u)+\cdots+a_L(x,u)} - x\sum_{j=0}^{L} P_j(a_0(x, u), \ldots, a_L(x, u), 1)). \end{aligned}$$

This system still fits our analytical framework. The advantage is that, for example, the *minimal* system of equations for counting stars in trees on page 24 now fits this modified system.

The idea for constructing $\mathscr{A}$ will be to create initially a certain family of tree classes $\mathscr{S} = \{t_1, \ldots, t_n\}$, not necessarily building a partition of p. Each of these classes will be defined as the class of all trees in p which 'start' in a certain way, or with other words, which match a certain tree t_i' at the root, just as was the case for the a_i in the main part of this paper. By abuse of notation we will usually write t_i instead of t_i' for this tree. Let $J = \{1, \ldots, n\}$ and $t_i^c = p \setminus t_i$. Now, by collecting in $\mathscr{A}$ all different, non-empty classes of the form

$$a_I = \bigcap_{i \in I} t_i \cap \bigcap_{i \in J \setminus I} t_i^c, \quad I \subseteq J \tag{A.1}$$

we will obtain a partition $\mathscr{A}$ of p. This partition will have a recursive description by construction, see the algorithms below. Furthermore, if $\mathscr{S}$ is sufficiently rich, this partition will allow an unambiguous definition of $K(l_0, \ldots, l_L)$.

We now make some considerations about the properties that $\mathscr{S}$ should possess to make sure that $\mathscr{A}$ will allow an unambiguous definition of $K(l_0, \ldots, l_L)$. Let b be a subclass of p. For each tree $T \in p$ we can determine the number $k(T)$ of pattern occurrences at the root of T. Let $k(b) = \{\, k(T) : T \in b \,\}$. Because the patterns have finitely many nodes and because in each internal node the degree is fixed and the root has to be part of the match, there are only finitely many ways for a pattern match. Thus the set $k(b)$ will be finite and non-empty. Now let a_I be defined by equation (A.1) (and non-empty). Now it holds that

$$k(a_I) \subseteq \bigcap_{i \in I} k(t_i) \cap \bigcap_{i \in J \setminus I} k(t_i^c) \tag{A.2}$$

because a tree T in a_I is by definition in t_i, $i \in I$ and t_i^c, $i \in J \setminus I$, and thus the number of pattern occurrences at the root is constrained by $k(t_i)$, $i \in I$ and $k(t_i^c)$, $i \in J \setminus I$. If $\mathscr{S} = \{t_1, \ldots, t_n\}$ is sufficiently rich, then $k(a_I)$ will only consist of a single number. This will be the case if, for each $m \in \mathbb{N}$, the family $\mathscr{S}$ contains all classes of trees 'starting' with all possible arrangements of m overlapping patterns. Indeed, if we have, for example, for a certain tree class t_i that $k(t_i) = \{r, r+1\}$, then there will be another tree class t_j, which is a subclass of t_i with $k(t_j) = \{r+1\}$. Now the intersections $b = t_i \cap t_j^c$ and $c = t_i \cap t_j$ will yield tree classes with a singleton $k(.)$, namely $k(b) = \{r\}$ and $k(c) = \{r+1\}$.

For example, consider a pattern which consists of a node of degree 2 attached to a node of degree 3. The corresponding planted patterns are shown in Figure 8. Now let $\mathscr{S}$ consist of the three classes t_1, t_2, t_3, shown in the centre of Figure 8. We have $k(t_1) = \{1\}$, because the left planted pattern surely matches and the other does not, $k(t_2) = \{1, 2\}$, because the left planted pattern does not match and the right one matches at least once, but possibly twice. Here $k(t_3) = \{2\}$, because the left pattern does not match and the right one surely matches twice. We see that the only non-empty intersections of the form (A.1) are $a = t_1 \cap t_2^c \cap t_3^c$, $b = t_1^c \cap t_2 \cap t_3^c$ and $c = t_1^c \cap t_2 \cap t_3$. We obtain $k(a) = k(b) = \{1\}$ and $k(c) = \{2\}$, which are all singletons. Because we also need a recursive description of

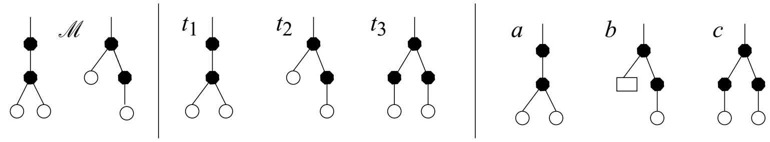

Figure 8. On the left: planted patterns. Centre: classes t_i. Right: classes $\{a, b, c\}$. The white box here means a node of out-degree different from 1. Note that this does not correspond to the output of the algorithms of this appendix.

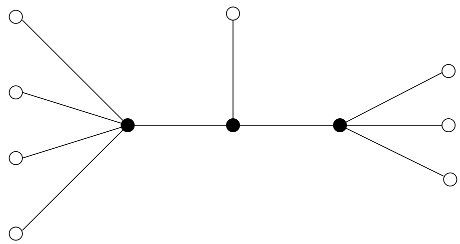

Figure 9. Example pattern $\mathscr{M}$.

the final partition $\mathscr{A}$, we will construct some additional tree classes t_i. As the partition becomes finer when dealing with more classes t_i, it is clear that k remains well-defined.

On the other hand we do not have to associate a unique number to $k(a_I)$, only to $K(l_0, \ldots, l_L)$. Therefore we can slightly reduce the family $\mathscr{S} = \{t_1, \ldots, t_n\}$. In the algorithm below this reduction of $\mathscr{S}$ corresponds to considering only proper subtrees of the trees $q \in \mathscr{Q}$ (q itself is excluded).

A coarse-grained description of an algorithm now follows.

(1) Calculate the set $\mathscr{U}$ of all planar embeddings of all planted patterns deducible from the pattern $\mathscr{M}$.
(2) Consider the planted planar trees issue of step (1) as planar tree classes and take all possible intersections of any number of those classes. Now take the implied non-planar general tree structure of each class and collect these non-planar planted trees in the set $\mathscr{Q}$.
(3) Create a family $\mathscr{S} = \{t_1, \ldots, t_n\}$ for the forest of planted subtrees of trees $q \in \mathscr{Q}$, excluding the trees q themselves, where each t_j has a recursive description in $t_0, t_1, \ldots, t_{j-1}$ and where t_0 denotes a leaf.
(4) Now interpret t_0 as the class of all trees p and interpret the trees $t_i \in \mathscr{S}$ as non-planar tree classes. Construct a partition $\mathscr{A} = \{a_0, \ldots, a_L\}$ of the class of all planted trees p together with a recursive description (compare with (A.1)).
(5) Calculate for each term in the recursive description the number $K(l_0, \ldots, l_L)$ of additional pattern occurrences and deduce a system of equations for the generating functions $a_j(x, u)$ of the classes a_j.

Before giving more detailed algorithms, we give an example. Consider the pattern of Figure 9.

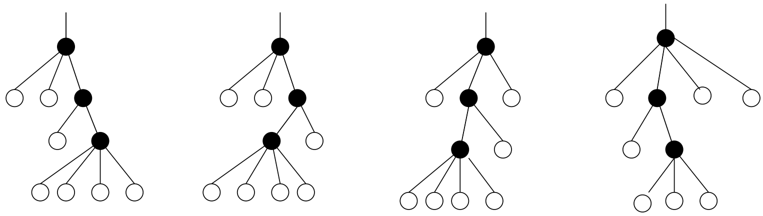

Figure 10. Some of in total 16 planted planar embeddings $\mathcal{U}$.

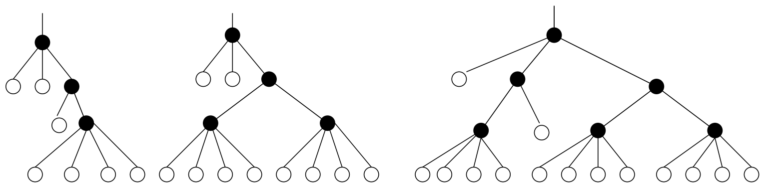

Figure 11. Some of in total 24 non-planar trees of $\mathcal{Q}$.

With the procedure of the main part of the article we would end up with more than 1000 classes, yielding a system of equations with the same number of equations. However, by using the following refined algorithm we only need 8 classes.

In the first step we create all planar embeddings of the corresponding planted pattern (trees τ_1, τ_2, τ_3 of Figure 14). This yields $3 \cdot 2 + 2 + 4 \cdot 2 = 16$ planar trees of which some are shown in Figure 10.

We now consider these structures as planar tree classes and additionally construct tree classes by taking all possible intersections of any number of the classes issued from step 1. Then, we take the non-planar implied tree structure of each planar class and collect these trees in $\mathcal{Q}$. We end up with 24 different trees: 9 that stem from τ_1, 1 from τ_2, and 14 from τ_3. Some of them are shown in Figure 11.

For all proper subtrees for each tree in $\mathcal{Q}$ we now construct a recursive description. For example, for the leftmost tree of Figure 11 we first consider the subtree consisting of a node with four leaves. We denote this class by $t_4 = xt_0^4$. (Here we use the following structural notation: x denotes a root node, t_0 a leaf and xt_0^4 denotes a root to which are attached 4 leaves.) The next subtree is a root of out-degree 2 to which a subtree of type t_4 is attached. We denote this with $t_5 = xt_0t_4$. Figure 12 shows all 6 trees we end up with. Observe in our example that the collection of subtrees at the root extracted from the 24 trees in $\mathcal{Q}$ consists of only 6 trees.

Their recursive descriptions are given by

$$t_1 = xt_0^3, \qquad t_2 = xt_0t_1, \qquad t_3 = xt_1^2, \qquad t_4 = xt_0^4, \qquad t_5 = xt_0t_4, \qquad t_6 = xt_4^2. \quad \text{(A.3)}$$

We now interpret t_0 in (A.3) as the class of all planted trees p. The other t_i are also interpreted as tree classes. For example, t_1 is the class of all trees with root out-degree 3. We now construct a partition based on these classes and their recursive description of (A.3). We obtain the classes of Figure 13.

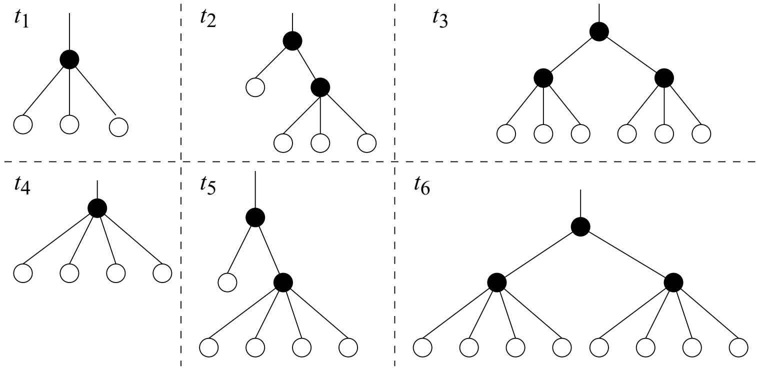

Figure 12. Non-planar trees t_i which possess a recursive description.

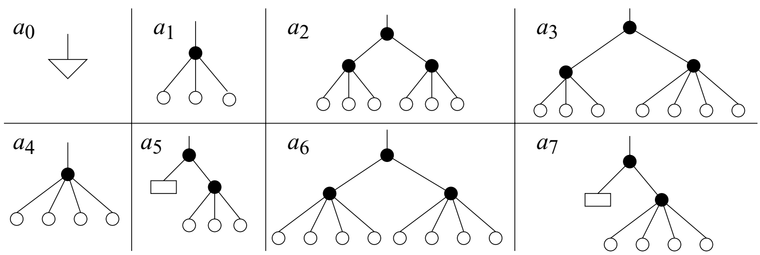

Figure 13. Non-planar partition classes. The white box means 'not out-degree 3 or 4' and the white triangle means 'anything that is not contained in the other classes'.

Their recursive description is given by

$$\begin{aligned}
a_0 &= 7p \setminus \bigoplus_{i=1}^{7} a_i = x \oplus x \bigoplus_{i=0}^{7} a_i \oplus x(a_0 \oplus a_2 \oplus a_3 \oplus a_5 \oplus a_6 \oplus a_7)^2 \oplus x \bigoplus_{n=5}^{\infty} \left(\bigoplus_{i=0}^{7} a_i \right)^n, \\
a_1 &= xp^3, \\
a_2 &= xa_1^2, \\
a_3 &= xa_1a_4, \\
a_4 &= xp^4, \\
a_5 &= x(a_0 \oplus a_2 \oplus a_3 \oplus a_5 \oplus a_6 \oplus a_7)a_1, \\
a_6 &= xa_4^2, \\
a_7 &= x(a_0 \oplus a_2 \oplus a_3 \oplus a_5 \oplus a_6 \oplus a_7)a_4.
\end{aligned} \tag{A.4}$$

The last step consists of determining the number of additional occurrences $K(l_0, \ldots, l_7)$ for each term in the recursive description (A.4) and translating (A.4) in a system of equations for the generating functions $a_j(x, u) = a_j$. As an example we consider the equation for a_1. Class a_1 consists of the trees of root out-degree 3. We get no additional occurrences of the pattern if we attach a tree of class a_0, a_1, a_2, a_4 or a_5 to such a root, we get one additional occurrence for each tree of class a_3 or a_7 and we have two additional occurrences for each tree of class a_6 attached to the root. This yields the equation for $a_1(x,u)$ below. Altogether we obtain

$$a_0 = x + x\sum_{i=0}^{7} a_i + \frac{1}{2}x(a_0+a_2+a_3+a_5+a_6+a_7)^2 + x\sum_{n\geqslant 5}\frac{1}{n!}\left(\sum_{i=0}^{7} a_i\right)^n,$$

$$a_1 = \frac{1}{3!}x(a_0+a_1+a_2+a_4+a_5+(a_3+a_7)u+a_6u^2)^3,$$

$$a_2 = \frac{1}{2}xa_1^2,$$

$$a_3 = xa_1a_4u,$$

$$a_4 = \frac{1}{4!}x(a_0+a_1+a_4+a_6+a_7+(a_3+a_5)u+a_6u^2)^4,$$

$$a_5 = x(a_0+a_2+a_3+a_5+a_6+a_7)a_1,$$

$$a_6 = \frac{1}{2}xa_4^2,$$

$$a_7 = x(a_0+a_2+a_3+a_5+a_6+a_7)a_4.$$

We can now calculate μ. We get $\mu = \frac{256-43e}{8e^3} = 0.865759040\ldots$. The computation of σ^2 was not feasible, because of memory problems.[5]

A.1. Planar embedding algorithm: GeneralToPlanar

Input. A general planted tree τ.

Output. The set $\mathscr{U}$ of planted planar trees π that share τ as their implied general tree structure.

Algorithm.

(1) Write τ in the form $x\tau_1\cdots\tau_k$, that is, let k be the root out-degree of τ and $\tau_1, \ldots, \tau_k$ be the children at the root.
(2) For each i between 1 and k, recursively compute $P_i = \text{GeneralToPlanar}(\tau_i)$.
(3) Construct and return the set of planar trees $x\pi_{\sigma(1)}\cdots\pi_{\sigma(k)}$ over all choices of $\pi_i \in P_i$ and over all permutations σ of $\{1,\ldots,k\}$.

A.2. Tree class intersection algorithm

Input. A set of planted planar trees $\mathscr{U}$.

Output. The set $\mathscr{Q}$ of non-planar planted trees which are obtained by intersecting planar tree classes based on $\mathscr{U}$ and collecting the non-planar tree structures of the resulting planar tree classes.

Algorithm.

(1) For each i between 1 and $|\mathscr{U}|$, consider all i-tuples of different trees $\pi_1,\ldots,\pi_i \in \mathscr{U}$ and determine for each i-tuple if $s = \pi_1 \cap \cdots \cap \pi_i$ may be interpreted as a non-empty tree

[5] The actual computation uses polynomial expressions with more than 200,000 terms. We used Maple 9.5, which used up GB of memory and a very large part of the 1 GB swap.

class. In that case, let s' be the implied non-planar tree structure of s and add s' to the set $\mathscr{Q}$.

A.3. DAGification algorithm

We construct a recursive description for the forest of planted subtrees for each tree in a given set of planted trees. Here we do not consider the tree itself as a subtree of itself. This calculation is reminiscent of the DAGification process of computer science (see, *e.g.*, [1]), which aims at compacting an expression tree by sharing repeated subexpressions. However, if we interpret those subtrees as classes, the intersection of two classes need not be empty.

Input. A set of planted trees $\mathscr{Q}$.

Output. A number m and a recursive description of the forest of planted subtrees $\mathscr{S} = \{t_1, \ldots, t_m\}$ of the trees of $\mathscr{Q}$, of the form

$$t_i = x t_{\lambda_1^{(i)}} \cdots t_{\lambda_{r_i}^{(i)}} \quad (r_i \in \mathbb{N}) \quad \text{for } 1 \leqslant i \leqslant m$$

with the constraint $\lambda_j^{(i)} < i$ for all i and j.

Algorithm.

(Initialisation) Introduce the exceptional type t_0 to denote the planted tree consisting of a single node (in other words, a leaf) and set m to 1.

(Main loop) For all planted trees of $\mathscr{U}$ perform a depth-first traversal of the tree, starting from the planted root; during this recursive calculation, at each node n:

(1) if the node is a leaf, return the type t_0
(2) else, recursively determine the type associated with each child of n
(3) if n is not the planted root of the tree, write the subtree rooted at n as a (commutative) product $\pi = x t_{\lambda_1} \cdots t_{\lambda_r}$ of the types obtained in the previous step
(4) look up the uniquification table to check whether this product has already been assigned a type t_i
(5) if not existent, increment m, create a new type t_m, remember its definition $t_m = \pi$, and assign t_m to the product π in the uniquification table
(6) return the type t_i if it was found by lookup, otherwise return t_m

(Conclusion) Return m and the sequence of definitions of the form $t_i = \pi$, for $i = 1, 2, \ldots, m$.

A.4. Disambiguating algorithm

The idea of the algorithm below is to consider each class of trees, t_i, in turn, introducing its defining equation

$$t_i = x t_{\lambda_1^{(i)}} \cdots t_{\lambda_{r_i}^{(i)}} \quad (r \in \mathbb{N})$$

into the calculation, while maintaining (and refining) a partition

$$p = a_0 \oplus \cdots \oplus a_L$$

of the total class of planted trees. To be able to do so, it is crucial that the recursive equation for t_i refers to classes t_j with $j < i$ only, starting with the special class $t_0 = p$, the full class of planted trees.

At any stage in the algorithm, the class of r-ary trees is given as the disjoint union of Cartesian products

$$\bigoplus_{\lambda \in \Lambda} x t_{\lambda_1} \cdots t_{\lambda_r} \quad \text{where} \quad \Lambda = \{\, \lambda : \ell(\lambda) = r,\ 0 \leqslant \lambda_j \leqslant L \,\},$$

where $\ell(\lambda)$ denotes the number of components in the tuple λ. In the process of the algorithm below, each class t_i gets represented in a 'polynomial' form as above, summed over a subset Λ of the set of integer sequences $\lambda = (\lambda_1, \ldots, \lambda_r)$ of a given length r. Computing intersections and differences of classes means merely computing intersections and differences of the Λ in their representations, because of the recursive structure of the input and of the algorithm itself.

Input.

- A family $\mathscr{S} = \{t_1, \ldots, t_m\}$ of classes of trees with recursive descriptions of the form

$$t_i = x t_{\lambda_1^{(i)}} \cdots t_{\lambda_r^{(i)}} \quad (r = \ell(\lambda^{(i)})) \quad \text{for} \quad 1 \leqslant i \leqslant m$$

  with the constraint $\lambda_j^{(i)} < i$ for all i and j.

Output.

- An integer L implying a partition

$$p = a_0 \oplus \cdots \oplus a_L$$

  .

- A representation of each t_i of the form

$$t_i = \bigoplus_{j \in I_i} a_j \quad \text{for} \quad 0 \leqslant i \leqslant m \quad \text{and} \quad I_i \subseteq \{0, \ldots, L\}.$$

- A recursive description of the a_i of the form

$$a_i = \bigoplus_{\lambda \in \Lambda_i} x a_{\lambda_1} \cdots a_{\lambda_{\ell(\lambda)}} \quad \text{for} \quad 1 \leqslant i \leqslant L,$$

  a_0 being implicitly described as $p \setminus (a_1 \oplus \cdots \oplus a_L)$.

Algorithm.

(Initialisation) Start with the trivial partition $p = a_0$ for $L = 0$, the single representation $t_0 = a_0$, that is, $I_0 = \{0\}$.

(Main loop) For k from 1 to m do

(1) replace each t_i in the definition of t_k with its current representation in terms of the a_j, expand, and set s to the result, so as to get a representation of t_k of the form

$$s = \bigoplus_{\lambda \in \Lambda^{(s)}} x a_{\lambda_1} \cdots a_{\lambda_{\ell(\lambda)}} \quad \text{for some} \quad \Lambda^{(s)}$$

(2) for i from 1 to L while $s \neq \emptyset$ do

(a) set b to $a_i \cap s$ by setting $\Lambda_\cap$ to $\Lambda_i \cap \Lambda^{(s)}$

(b) if $b \neq \emptyset$, then do

(i) set b' to $a_i \setminus s$

(ii) if $b' \neq \emptyset$, then

(A) create a new a_j with description b': increment n before setting a_L to b', that is, before setting Λ_L to $\Lambda_i \setminus \Lambda^{(s)}$

(B) split a_i into $a_i \oplus a_L$ in the representations of the t_j, that is, add n into each set I_j containing i

(C) split a_i into $a_i \oplus a_L$ in the descriptions of the a_j, b, and s, that is, for each sequence in each of the Λ_j, $\Lambda_\cap$, and $\Lambda^{(s)}$, add sequences with i replaced by L when the sequence involves i (if i occurs more than once, then replace i by i or L in all possible ways)

(D) set a_i to b by setting Λ_i to $\Lambda_\cap$

(iii) set s to $s \setminus b$, which is also $s \setminus a_i$, and update $\Lambda^{(s)}$ by setting it to $\Lambda^{(s)} \setminus \Lambda_i$

(3) if $s \neq \emptyset$, then

(a) create a new a_j with description s: increment L before setting a_L to s, that is, before setting Λ_L to $\Lambda^{(s)}$

(b) split a_0 into $a_0 \oplus a_L$ in the representations of the t_j, that is, add L into each set I_j containing 0

(c) split a_0 into $a_0 \oplus a_L$ in the descriptions of the a_j, that is, for each sequence in each of the Λ_j, add sequences with 0 replaced by n when the sequence involves 0 (if 0 occurs more than once, then replace 0 by 0 or L in all possible ways)

(4) represent t_k as the union of all those a_is that have contributed a non-empty b at step (2(b)) and of a_L if a new a_j was created at step (3(a)) that is, create the corresponding set I_k consisting of the contributing is, together with L if relevant

(Final step) Return L, the representations of the t_i for $1 \leqslant i \leqslant m$, the recursive descriptions of the a_i for $1 \leqslant i \leqslant L$.

We will explicitly show the stages through which the algorithm goes when running with the input (A.3). For readability, we will keep expressions in factored form.

$k = 1$: from $t_1 = xa_0^3$, we derive $t_1 = a_1$ and $a_1 = x(a_0 \oplus a_1)^3$.

$k = 2$: from $t_2 = x(a_0 \oplus a_1)a_1$, we derive $t_1 = a_1$, $t_2 = a_2$ and $a_1 = xp^3$, $a_2 = xpa_1$, where $p = a_0 \oplus a_1 \oplus a_2$.

$k = 3$: from $t_3 = xa_1^2$, we derive $t_1 = a_1$, $t_2 = a_2 \oplus a_3$, $t_3 = a_2$ and $a_1 = xp^3$, $a_2 = xa_1^2$, $a_3 = x(a_0 \oplus a_2 \oplus a_3)a_1$, where $p = a_0 \oplus a_1 \oplus a_2 \oplus a_3$.

$k = 4$: from $t_4 = x(a_0 \oplus a_1 \oplus a_2 \oplus a_3)^4$, we derive $t_1 = a_1$, $t_2 = a_2 \oplus a_3$, $t_3 = a_2$, $t_4 = a_4$ and $a_1 = xp^3$, $a_2 = xa_1^2$, $a_3 = x(a_0 \oplus a_2 \oplus a_3 \oplus a_4)a_1$, $a_4 = xp^4$, where $p = a_0 \oplus a_1 \oplus a_2 \oplus a_3 \oplus a_4$.

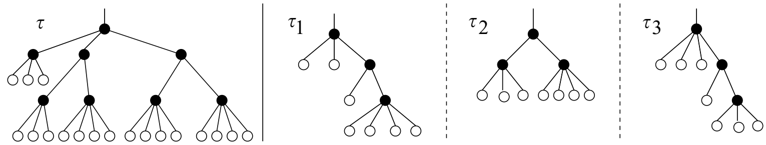

Figure 14. Input trees $\tau, \tau_1, \tau_2, \tau_3$.

$k = 5$: from $t_5 = x(a_0 \oplus a_1 \oplus a_2 \oplus a_3 \oplus a_4)a_4$, we derive $t_1 = a_1$, $t_2 = a_2 \oplus a_3 \oplus a_5$, $t_3 = a_2$, $t_4 = a_4$, $t_5 = a_3 \oplus a_6$ and $a_1 = xp^3$, $a_2 = xa_1^2$, $a_3 = xa_1a_4$, $a_4 = xp^4$, $a_5 = x(a_0 \oplus a_2 \oplus a_3 \oplus a_5 \oplus a_6)a_1$, $a_6 = x(a_0 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5 \oplus a_6)a_4$, where $p = a_0 \oplus a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5 \oplus a_6$.

$k = 6$: from $t_6 = xa_4^2$, we derive $t_1 = a_1$, $t_2 = a_2 \oplus a_3 \oplus a_5$, $t_3 = a_2$, $t_4 = a_4$, $t_5 = a_3 \oplus a_6 \oplus a_7$, $t_6 = a_6$ and $a_1 = xp^3$, $a_2 = xa_1^2$, $a_3 = xa_1a_4$, $a_4 = xp^4$, $a_5 = x(a_0 \oplus a_2 \oplus a_3 \oplus a_5 \oplus a_6 \oplus a_7)a_1$, $a_6 = xa_4^2$, $a_7 = x(a_0 \oplus a_2 \oplus a_3 \oplus a_5 \oplus a_6 \oplus a_7)a_4$, where $p = a_0 \oplus a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5 \oplus a_6 \oplus a_7$.

A.5. Calculation of $K(l_0, \ldots, l_L)$: CountRootOccurrences

Input. Non-planar planted trees τ, τ_1, ... , τ_k.

Output. The number of occurrences of any of the τ_i at the root of τ.

Algorithm.

(1) Fix one element π' from GeneralToPlanar(τ) (see algorithm A.1).
(2) For each i between 1 and k, compute P_i = GeneralToPlanar(τ).
(3) Count and return the number of pairs (π_i, π') such that π_i is element of P_i and π_i occurs at the root of π'.

As an example we calculate $K(0, 1, 0, 1, 0, 0, 1, 0)$. This corresponds to calculating the number of additional occurrences in the class $xa_1a_3a_6$. The input trees $\tau, \tau_1, \tau_2, \tau_3$ are shown in Figure 14. Here τ corresponds to the class $xa_1a_3a_6$ and τ_1, τ_2, τ_3 correspond to the three possible ways of planting the example pattern.

We take as fixed planar embedding π' of τ the embedding of Figure 14. We now iterate over the different planar embeddings π_1 of τ_1 (6 of them), π_2 of τ_2 (2 of them), and π_3 of τ_3 (8 of them), and determine for each π_i ($i \in \{1, 2, 3\}$) whether it occurs at the root of π'. Consider, for example, the four embeddings shown in Figure 10 (three embeddings of τ_1, one embedding of τ_3). The leftmost embedding matches π', the one next to it as well. The third one does not match π', because the node with out-degree four is in the wrong position. The rightmost embedding clearly does not match either. By considering all embeddings and counting the matches we get $k = K(0, 1, 0, 1, 0, 0, 1, 0) = 3$.

The algorithm calculates the correct value of k, because the partition consisting of the classes a_i is sufficiently fine. From this it follows that every match above of a planar embedding really gives rise to exactly one additional pattern occurrence. See the considerations made at the beginning of this appendix.

By now the transformation to a systems of equations is easy. We get the terms by replacing a term $xa_{j_1} \cdots a_{j_s}$ in the recursive description of a_j by a term $xy_{j_1} \cdots y_{j_s}u^{K(l_0,\ldots,l_L)}/l_0! \ldots l_L!$. Here it is assumed that terms that represent the same tree classes (like xa_1a_2 and

xa_2a_1) are identified before. It is clear that there are only finitely many terms for which $K(l_0, \dots, l_L)$ might be non-zero *a priori*.

Appendix B: Asymptotics of analytic systems

The following theorem is a slightly modified version of the main theorem from [4]. We denote the transpose of a vector v by v^T. Let $\mathbf{F}(x, \mathbf{y}, \mathbf{u}) = (F_1(x, \mathbf{y}, \mathbf{u}), \dots, F_N(x, \mathbf{y}, \mathbf{u}))^T$ be a column vector of functions $F_j(x, \mathbf{y}, \mathbf{u})$, $1 \leqslant j \leqslant N$, with complex variables x, $\mathbf{y} = (y_1, \dots, y_N)^T$, $\mathbf{u} = (u_1, \dots, u_k)^T$ which are analytic around 0 and satisfy $F_j(0, \mathbf{0}, \mathbf{0}) = 0$ for $1 \leqslant j \leqslant N$. We are interested in the analytic solution $\mathbf{y} = \mathbf{y}(x, \mathbf{u}) = (y_1(x, \mathbf{u}), \dots, y_N(x, \mathbf{u}))^T$ of the functional equation

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{u}) \tag{B.1}$$

with $\mathbf{y}(0, \mathbf{0}) = \mathbf{0}$, *i.e.*, we demand that the (unknown) functions $y_j = y_j(x, \mathbf{u})$, $1 \leqslant j \leqslant N$, satisfy the system of functional equations

$$\begin{aligned}
y_1 &= F_1(x, y_1, y_2, \dots, y_N, \mathbf{u}), \\
y_2 &= F_2(x, y_1, y_2, \dots, y_N, \mathbf{u}), \\
&\vdots \\
y_N &= F_N(x, y_1, y_2, \dots, y_N, \mathbf{u}).
\end{aligned}$$

It is convenient to define the notion of a dependency (di)graph $G_\mathbf{F} = (V, E)$ for such a system of functional equations $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{u})$. The vertices $V = \{y_1, y_2, \dots, y_N\}$ are just the unknown functions and an ordered pair (y_i, y_j) is contained in the edge set E if and only if $F_i(x, \mathbf{y}, \mathbf{u})$ really depends on y_j.

If the functions $F_j(x, \mathbf{y}, \mathbf{u})$ have non-negative Taylor coefficients then it is easy to see that the solutions $y_j(x, \mathbf{u})$ have the same property. (One only has to solve the system iteratively by setting $\mathbf{y}_0(x, \mathbf{u}) = 0$ and $\mathbf{y}_{i+1}(x, \mathbf{u}) = \mathbf{F}(x, \mathbf{y}_i(x, \mathbf{u}), \mathbf{u})$ for $i \geqslant 0$. The limit $\mathbf{y}(x, \mathbf{u}) = \lim_{i \to \infty} \mathbf{y}_i(x, \mathbf{u})$ is the (unique) solution of the system above.)

Now suppose that $G(x, \mathbf{y}, \mathbf{u})$ is another analytic function with non-negative Taylor coefficients. Then $G(x, \mathbf{y}(x, \mathbf{u}), \mathbf{u})$ has a power series expansion

$$G(x, \mathbf{y}(x, \mathbf{u}), \mathbf{u}) = \sum_{n, \mathbf{m}} c_{n,\mathbf{m}} x^n \mathbf{u}^\mathbf{m}$$

with non-negative coefficients $c_{n,\mathbf{m}}$. In fact, we assume that for every $n \geqslant n_0$ there exists $\mathbf{m}$ such that $c_{n,\mathbf{m}} > 0$.

Let $\mathbf{X}_n$ ($n \geqslant n_0$) denote an N-dimensional discrete random vector with

$$\mathbf{Pr}[\mathbf{X}_n = \mathbf{m}] := \frac{c_{n,\mathbf{m}}}{c_n}, \tag{B.2}$$

where

$$c_n = \sum_{\mathbf{m}} c_{n,\mathbf{m}}$$

are the coefficients of

$$G(x, \mathbf{y}(x, \mathbf{1}), \mathbf{1}) = \sum_{n \geqslant 0} c_n x^n.$$

The following theorem shows that (under suitable analyticity conditions) $\mathbf{X}_n$ has a Gaussian limiting distribution.

Theorem B.1. *Let* $\mathbf{F}(x, \mathbf{y}, \mathbf{u}) = (F_1(x, \mathbf{y}, \mathbf{u}), \dots, F_N(x, \mathbf{y}, \mathbf{u}))^{\mathrm{T}}$ *be functions analytic around* $x = 0$, $\mathbf{y} = (y_1, \dots, y_N)^{\mathrm{T}} = \mathbf{0}$, $\mathbf{u} = (u_1, \dots, u_k)^{\mathrm{T}} = \mathbf{0}$, *whose Taylor coefficients are all non-negative, such that* $\mathbf{F}(0, \mathbf{y}, \mathbf{u}) = \mathbf{0}$, $\mathbf{F}(x, \mathbf{0}, \mathbf{u}) \neq \mathbf{0}$, $\mathbf{F}_x(x, \mathbf{y}, \mathbf{u}) \neq \mathbf{0}$, *and such that there exists* j *with* $\mathbf{F}_{y_j y_j}(x, \mathbf{y}, \mathbf{u}) \neq \mathbf{0}$. *Furthermore, assume that the region of convergence of* $\mathbf{F}$ *is large enough that there exists a non-negative solution* $x = x_0$, $\mathbf{y} = \mathbf{y}_0$ *of the system of equations*

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{1}),$$
$$0 = \det(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x, \mathbf{y}, \mathbf{1})),$$

inside it. Let

$$\mathbf{y} = \mathbf{y}(x, \mathbf{u}) = (y_1(x, \mathbf{u}), \dots, y_N(x, \mathbf{u}))^{\mathrm{T}}$$

denote the analytic solutions of the system

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{u}) \tag{B.3}$$

with $\mathbf{y}(0, \mathbf{u}) = \mathbf{0}$ *and assume that* $d_{n,j} > 0$ $(1 \leqslant j \leqslant N)$ *for* $n \geqslant n_1$, *where* $y_j(x, \mathbf{1}) = \sum_{n \geqslant 0} d_{n,j} x^n$. *Moreover, let* $G(x, \mathbf{y}, \mathbf{u})$ *denote an analytic function with non-negative Taylor coefficients such that the point* $(x_0, \mathbf{y}(x_0, \mathbf{1}), \mathbf{1})$ *is contained in the region of convergence. Finally, let random vectors* $\mathbf{X}_n$ $(n \geqslant n_0)$ *be defined by* (B.2).

If the dependency graph $G_{\mathbf{F}} = (V, E)$ *of the system* (B.3) *in the unknown functions* $y_1(x, \mathbf{u})$, $\dots$, $y_N(x, \mathbf{u})$ *is strongly connected then the sequence of random vectors* $\mathbf{X}_n$ *admits a Gaussian limiting distribution with mean value*

$$\mathbf{E}\,\mathbf{X}_n = \boldsymbol{\mu}\, n + O(1) \quad (n \to \infty)$$

and covariance matrix

$$\mathbf{Cov}(\mathbf{X}_n, \mathbf{X}_n) = \Sigma\, n + O(1) \quad (n \to \infty).$$

The row vector $\boldsymbol{\mu}$ *is given by*

$$\boldsymbol{\mu} = -\frac{x_{\mathbf{u}}(\mathbf{1})}{x(\mathbf{1})},$$

and the matrix Σ *by*

$$\Sigma = -\frac{\mathbf{x}_{\mathbf{uu}}(\mathbf{1})}{\mathbf{x}(\mathbf{1})} + \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{\mu} + \mathrm{diag}(\boldsymbol{\mu}), \tag{B.4}$$

where $x = x(\mathbf{u})$ (*and* $\mathbf{y} = \mathbf{y}(\mathbf{u}) = \mathbf{y}(x(\mathbf{u}), \mathbf{u})$) *is the solution of the* (*extended*) *system*

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{u}), \tag{B.5}$$
$$0 = \det(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x, \mathbf{y}, \mathbf{u})). \tag{B.6}$$

The proof of Theorem B.1 is exactly the same as that given in [4]. The main observation is that the assumptions above show that the solutions $y_j(x, \mathbf{u})$ admit a local representation of the form

$$y_j(x,\mathbf{u}) = g_j(x,\mathbf{u}) - h_j(x,\mathbf{u})\sqrt{1-\frac{x}{x(\mathbf{u})}},$$

(where $\mathbf{u}$ is close to $\mathbf{1}$ and x close to $x_0 = x(\mathbf{1})$). The assumption that the dependency graph is strongly connected ensures that the location of the singularity of all functions $y_j(x,\mathbf{u})$ is determined by the common function $x(\mathbf{u})$. Thus, we get the same property for $G(x,\mathbf{y}(x,\mathbf{u}),\mathbf{u})$:

$$G(x,\mathbf{y}(x,\mathbf{u}),\mathbf{u}) = g(x,\mathbf{u}) - h(x,\mathbf{u})\sqrt{1-\frac{x}{x(\mathbf{u})}}. \tag{B.7}$$

It is then well known (see [2, 3]) that a square root singularity plus some minor conditions implies asymptotic normality of the coefficients (in the sense introduced above) with mean and covariance expressed in terms of derivatives of $x(\mathbf{u})$. Note, for example, that the assumption $d_{n,j} > 0$ for $n \geqslant n_1$ ensures that $c_n > 0$ for sufficiently large n and from this follows that $x_0 = x(\mathbf{1})$ is the only singularity on the radius of convergence of $G(x,\mathbf{y}(x,\mathbf{1}),\mathbf{1})$.

In what follows we comment on the evaluation of $\boldsymbol{\mu}$ and Σ. The problem is to extract the derivatives of $x(\mathbf{u})$. The function $x(\mathbf{u})$ is the solution of the system (B.5)–(B.6) and is exactly the location of the singularity of the mapping $x \mapsto \mathbf{y}(x,\mathbf{u})$ when $\mathbf{u}$ is fixed (and close to $\mathbf{1}$).

Let $x(\mathbf{u})$ and $\mathbf{y}(\mathbf{u}) = \mathbf{y}(x(\mathbf{u}),\mathbf{u})$ denote the solutions of (B.5)–(B.6). Then we have

$$\mathbf{y}(\mathbf{u}) = \mathbf{F}(x(\mathbf{u}),\mathbf{y}(\mathbf{u}),\mathbf{u}). \tag{B.8}$$

Taking derivatives with respect to $\mathbf{u}$ we get

$$\mathbf{y}_{\mathbf{u}}(\mathbf{u}) = \mathbf{F}_x(x(\mathbf{u}),\mathbf{y}(\mathbf{u}),\mathbf{u})x_{\mathbf{u}}(\mathbf{u}) + \mathbf{F}_{\mathbf{y}}(x(\mathbf{u}),\mathbf{y}(\mathbf{u}),\mathbf{u})\mathbf{y}_{\mathbf{u}}(\mathbf{u}) + \mathbf{F}_{\mathbf{u}}(x(\mathbf{u}),\mathbf{y}(\mathbf{u}),\mathbf{u}), \tag{B.9}$$

where the three terms in $\mathbf{F}$ denote evaluations at $(x(\mathbf{u}),\mathbf{y}(\mathbf{u}),\mathbf{u})$ of the partial derivatives of $\mathbf{F}$, and where $x_{\mathbf{u}}$ and $\mathbf{y}_{\mathbf{u}}$ denote the Jacobian of x resp. $\mathbf{y}$ with respect to $\mathbf{u}$. In particular, for $\mathbf{u} = \mathbf{1}$ we have $x(\mathbf{1}) = x_0$ and $\mathbf{y}(\mathbf{1}) = \mathbf{y}_0$ and, of course,

$$\det(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x_0,\mathbf{y}_0,\mathbf{1})) = 0.$$

Since $\mathbf{F}_{\mathbf{y}}$ is a non-negative matrix and the dependency graph is strongly connected there is a unique Perron–Frobenius eigenvalue of multiplicity 1. Here this eigenvalue equals 1. Thus, $\mathbf{I} - \mathbf{F}_{\mathbf{y}}$ has rank $N-1$ and has (up to scaling) a unique positive left eigenvector $\mathbf{b}^{\mathrm{T}}$:

$$\mathbf{b}^{\mathrm{T}}(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x_0,\mathbf{y}_0,\mathbf{1})) = \mathbf{0}.$$

From (B.9) we obtain

$$(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x_0,\mathbf{y}_0,\mathbf{1}))\mathbf{y}_{\mathbf{u}}(\mathbf{1}) = \mathbf{F}_x(x_0,\mathbf{y}_0,\mathbf{1})x_{\mathbf{u}}(\mathbf{1}) + \mathbf{F}_{\mathbf{u}}(x_0,\mathbf{y}_0,\mathbf{1}).$$

By multiplying $\mathbf{b}^{\mathrm{T}}$ from the left we thus get

$$\mathbf{b}^{\mathrm{T}}\mathbf{F}_x(x_0,\mathbf{y}_0,\mathbf{1})x_{\mathbf{u}} + \mathbf{b}^{\mathrm{T}}\mathbf{F}_{\mathbf{u}}(x_0,\mathbf{y}_0,\mathbf{1}) = 0 \tag{B.10}$$

and consequently

$$\mu = \frac{1}{x_0}\frac{\mathbf{b}^{\mathrm{T}}\mathbf{F}_{\mathbf{u}}(x_0,\mathbf{y}_0,\mathbf{1})}{\mathbf{b}^{\mathrm{T}}\mathbf{F}_x(x_0,\mathbf{y}_0,\mathbf{1})}.$$

The derivation of Σ is more involved. We first define $\mathbf{b}(x, \mathbf{y}, \mathbf{u})$ as the (generalized) vector product[6] of the $N-1$ last columns of the matrix $\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x, \mathbf{y}, \mathbf{u})$. Observe that

$$D(x, \mathbf{y}, \mathbf{u}) := \left(\mathbf{b}^T(x, \mathbf{y}, \mathbf{u})(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x, \mathbf{y}, \mathbf{u}))\right)_1 = \det(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x, \mathbf{y}, \mathbf{u})).$$

In particular, we have

$$D(x(\mathbf{u}), \mathbf{y}(\mathbf{u}), \mathbf{u}) = 0.$$

Then from

$$\begin{aligned}(\mathbf{I} - \mathbf{F}_{\mathbf{y}})\mathbf{y}_{\mathbf{u}} &= \mathbf{F}_x x_{\mathbf{u}} + \mathbf{F}_{\mathbf{u}}, \\ -D_{\mathbf{y}}\mathbf{y}_{\mathbf{u}} &= D_x x_{\mathbf{u}} + D_{\mathbf{u}}\end{aligned} \qquad \text{(B.11)}$$

we can calculate $\mathbf{y}_{\mathbf{u}}$. (The first system has rank $N-1$: this means that we can skip the first equation. This reduced system is then completed to a regular system by appending the second equation (B.11).)

We now set

$$\begin{aligned}d_1(\mathbf{u}) &= d_1(x(\mathbf{u}), \mathbf{y}(\mathbf{u}), \mathbf{u}) = \mathbf{b}(x(\mathbf{u}), \mathbf{y}(\mathbf{u}), \mathbf{u})^{\mathrm{T}} \mathbf{F}_x(x(\mathbf{u}), \mathbf{y}(\mathbf{u}), \mathbf{u}), \\ \mathbf{d}_2(\mathbf{u}) &= \mathbf{d}_2(x(\mathbf{u}), \mathbf{y}(\mathbf{u}), \mathbf{u}) = \mathbf{b}(x(\mathbf{u}), \mathbf{y}(\mathbf{u}), \mathbf{u})^{\mathrm{T}} \mathbf{F}_{\mathbf{u}}(x(\mathbf{u}), \mathbf{y}(\mathbf{u}), \mathbf{u}).\end{aligned}$$

By differentiating equation (B.10) we get

$$x_{\mathbf{uu}}(\mathbf{u}) = -\frac{(d_{1x}x_{\mathbf{u}} + d_{1\mathbf{y}}\mathbf{y}_{\mathbf{u}} + d_{1\mathbf{u}})x_{\mathbf{u}} + (\mathbf{d}_{2x}x_{\mathbf{u}} + \mathbf{d}_{2\mathbf{y}}\mathbf{y}_{\mathbf{u}} + \mathbf{d}_{2\mathbf{u}})}{d_1}, \qquad \text{(B.12)}$$

where $d_{1x}, d_{1\mathbf{y}}, d_{1\mathbf{u}}, \mathbf{d}_{2x}, \mathbf{d}_{2\mathbf{y}}, \mathbf{d}_{2\mathbf{u}}$ denote the respective partial derivatives and where we omitted the dependence on $\mathbf{u}$. With the knowledge of $x_0, \mathbf{y}_0$ and $\mathbf{y}_{\mathbf{u}}(\mathbf{1})$ we can now evaluate $x_{\mathbf{uu}}$ at $\mathbf{u} = \mathbf{1}$ and we finally calculate Σ from (B.4).

Appendix C: Proof of Lemma 4.3

In this appendix we will prove Lemma 4.3 saying that the determinant $\det(\mathbf{I} - \mathbf{F}_{\mathbf{a}}(x, \mathbf{a}, 1))$ is given by

$$\det(\mathbf{I} - \mathbf{F}_{\mathbf{a}}(x, \mathbf{a}, 1)) = 1 - xe^{a_0 + a_1 + \cdots + a_L}.$$

We first observe that the sum of all rows of $\mathbf{I} - \mathbf{F}_{\mathbf{a}}(x, \mathbf{a}, 1)$ equals

$$\left(1 - xe^{a_0 + a_1 + \cdots + a_L}, 1 - xe^{a_0 + a_1 + \cdots + a_L}, \ldots, 1 - xe^{a_0 + a_1 + \cdots + a_L}\right),$$

compare with (4.9). Hence, we get

$$\det(\mathbf{I} - \mathbf{F}_{\mathbf{a}}(x, \mathbf{a}, 1)) = (1 - xe^{a_0 + a_1 + \ldots + a_L}) \det \mathbf{M}(x, \mathbf{a}),$$

where $\mathbf{M}(x, \mathbf{a})$ denotes the matrix $\mathbf{I} - \mathbf{F}_{\mathbf{a}}$ where we replace the first row by $(1, 1, \ldots, 1)$. Thus, it remains to prove that $\det \mathbf{M}(x, \mathbf{a}) = 1$.

For this purpose we have to be more explicit with the partition $\mathscr{A} = \{a_0, a_1, \ldots, a_L\}$. More precisely we construct $\mathscr{A}$ recursively from level to level. This procedure is similar

[6] More precisely, this is the wedge product combined with the Hodge duality.

to that of Proposition 4.1 but not the same. In order to make our arguments more transparent we restrict ourselves to 4 steps. Note that this procedure also provides a recursive description of the polynomials $P_j(\mathbf{a}, 1)$.

One starts with $\mathscr{A}_0 = \{d_0, d_1\}$, where $d_0 = a_0$ and $d_1 = p \setminus a_0$. This means that d_0 collects all trees where the root out-degree is not contained in D and d_1 those where it is contained in D. For example, if $D = \{2\}$ then the generating functions of this (trivial) partition are given by $d_1(x, 1) = xp(x)^2/2$ and by $d_0(x, 1) = p(x) - d_1(x, 1) = p(x) - xp(x)^2/2$.

Then we partition d_1 according to structure of the subtrees of the root, where we distinguish between the previous classes d_0 and d_1. We get $\mathscr{A}_1 = \{c_0, c_1, \ldots, c_m\}$, where $c_0 = d_0$ and $c_1 \oplus \ldots \oplus c_m = d_1$. In particular, if $D = \{2\}$ then $m = 3$, the class c_1 collects all trees with root out-degree 2 where both subtrees of the root are in class $a_0 = d_0$, c_2 collects all trees with root out-degree 2 where one subtree of the root is in class $a_0 = d_0$ and the other one in class d_1, and c_3 collects those trees where both subtrees of the root are in class d_1. The corresponding generating functions are given by $c_1(x, 1) = xd_0(x, 1)^2/2$, by $c_2(x, 1) = xd_0(x, 1)d_1(x, 1)$, and by $c_3(x, 1) = xd_1(x, 1)^2/2$. Of course, we also have $c_0(x, 1) = d_0(x, 1)$ and $c_1(x, 1) + c_2(x, 1) + c_3(x, 1) = d_1(x, 1)$.

In the same fashion we proceed further. We partition c_s ($1 \leqslant s \leqslant m$) according to the structure of the subtrees of the root (which are now taken from $\{c_1, \ldots, c_m\}$) and denote them by $\mathscr{A}_2 = \{b_0, b_1, \ldots, b_\ell\}$. Further we define sets C_s by $c_s = \bigoplus_{r \in C_s} b_r$. If $D = \{2\}$ then $b_0 = c_0$, $b_1 = c_1$, c_2 is divided into three parts, and c_3 is divided into 6 parts: $C_1 = \{1\}$, $C_2 = \{2, 3, 4\}$, $C_3 = \{5, 6, 7, 8, 9, 10\}$.[7]

Finally, we partition b_j ($j \geqslant 1$) according to the structure of the subtrees of the root that are taken from the b_i and denote them by $\mathscr{A} = \{a_0, a_1, \ldots, a_L\}$. As in the previous step we define sets B_r by $b_r = \bigoplus_{j \in B_r} a_j$. In general we have to iterate this procedure until a certain level and get almost the same partition as in the proof of Proposition 3.1. The only difference is that at the lowest level we only distinguish between nodes with degree in D and degree not in D. However, this is no real restriction as we can extend the partition above with an additional level and we will have a well-defined number of additional occurrences for each class. We again obtain a partition which fits Proposition 3.1.

We recall that this recursive procedure directly provides a recursive description of the system of functional equations. In particular, we have

$$a_j(x, 1) = x\, P_j(a_0(x, 1), a_1(x, 1), \ldots, a_L(x, 1), 1),$$

where $P_j(\cdot, 1)$ can actually be written as a polynomial in $b_0, b_1, \ldots, b_\ell$.

Next

$$b_r(x, 1) = x\, Q_r(b_0(x, 1), b_1(x, 1), \ldots, b_\ell(x, 1), 1),$$

where $Q_r(\cdot, 1)$ can be written as a polynomial in $c_0, c_1, \ldots, c_m$. Further,

$$Q_r = \sum_{j \in B_r} P_j.$$

In other words, the sum $\sum_{j \in B_r} P_j$ can be written as a polynomial in c_r.

[7] By the way this leads to the partition that is used in the proof of Theorem 1.1 resp. of Proposition 3.1.

Finally,

$$c_s(x,1) = x\,R_s(c_0(x,1), c_1(x,1), \ldots, c_m(x,1)),$$

where $R_s(\cdot,1)$ can be written as a polynomial in $d_0 = a_0$ and $d_1 = a_1 + \cdots + a_L$ and we have

$$R_s = \sum_{r\in C_s} Q_r.$$

Let $\mathbf{G}(x,\mathbf{a})$ denote the $L \times L$-submatrix of $\mathbf{F_a}$ where we omit the first row and column. Then $\mathbf{G}(x,\mathbf{a})$ has the following structure:

$$\mathbf{G}(x,\mathbf{a}) = \begin{pmatrix} G_{11} & \cdots & G_{1m} \\ \vdots & & \vdots \\ G_{m1} & \cdots & G_{mm} \end{pmatrix},$$

where

$$G_{s's''} = \left(B_{r'r''}\right)_{r'\in C_{s'}, r''\in C_{s''}}$$

and

$$B_{r'r''} = \left(xP_{i,a_j}\right)_{i\in B_{r'}, j\in B_{r''}}.$$

The condition that P_i can be written as a polynomial in b_j implies that $P_{i,a_{j_1}} = P_{i,a_{j_2}}$ for all $j_1, j_2 \in B_r$, that is, each row of $B_{r'r''}$ is either zero or all entries are the same.

Further, if we fix r' and sum over all rows $i \in B_{r'}$ then we get

$$\sum_{i\in B_{r'}} xP_{i,a_j} = xQ_{r',a_j}.$$

Since $Q_{r'}$ can be written as a polynomial in c_s $(0 \leqslant s \leqslant m)$ we have $Q_{r',a_{j_1}} = Q_{r',a_{j_2}}$ for all $j_1, j_2 \in \bar{C}_{s''}$, where we set $\bar{C}_s = \bigcup_{r\in C_s} B_r$.

Similarly if we fix s' and sum over all rows $i \in \bar{C}_{s'}$ then we get

$$\sum_{i\in \bar{C}_{s'}} xP_{i,a_j} = xR_{s',a_j}.$$

Since $R_{s'}$ can be written as a polynomial in $d_0 = a_0$ and $d_1 = a_1 + \cdots + a_L$ we have $R_{s',a_{j_1}} = R_{s',a_{j_2}}$ for all $1 \leqslant j_1, j_2 \leqslant L$.

Now we will calculate the determinant of the matrix

$$\mathbf{M}(x,\mathbf{a}) = \begin{pmatrix} 1 & 1\cdots & \cdots & \cdots 1 \\ 0 & \mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \mathbf{0} & \cdots & \mathbf{I} \end{pmatrix} - \begin{pmatrix} 0 & \mathbf{0} & \cdots & \mathbf{0} \\ \times & G_{11} & \cdots & G_{1m} \\ \vdots & \vdots & & \vdots \\ \times & G_{m1} & \cdots & G_{mm} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1\cdots & \cdots & \cdots 1 \\ \times & \mathbf{I} - G_{11} & \cdots & -G_{1m} \\ \vdots & \vdots & & \vdots \\ \times & -G_{m1} & \cdots & \mathbf{I} - G_{mm} \end{pmatrix}.$$

(By $\times$ we denote an entry whose value is immaterial.) We now perform the following row operations. For every $s' = 1, \ldots, m$ we substitute the first row of

$$\left(\times \; -G_{s'1} \; \cdots \; \mathbf{I} - G_{s's'} \; \cdots - G_{s'm}\right)$$

by the sum of the corresponding rows $i \in \bar{C}_{s'}$. Since $R_{s',a_{j_1}} = R_{s',a_{j_2}}$ for all $1 \leqslant j_1, j_2 \leqslant L$ this sum of the rows has the form

$$\left(\times \; -xR_{s',a} \cdots - xR_{s',a} \; \cdots \; 1 - xR_{s',a} \cdots 1 - xR_{s',a} \; \cdots \; -xR_{s',a} \cdots - xR_{s',a}\right).$$

We now add the very first row (which equals $(1, 1, \ldots, 1)$) $xR_{s',a}$ times to this row and obtain

$$\mathbf{w}_{s'} = \left(\times \;\; | \;\; 0 \cdots 0 \;\; | \;\; \cdots \;\; | \;\; 1 \cdots 1 \;\; | \;\; \cdots \;\; | \;\; 0 \cdots 0\right).$$

Next we fix s' and r' such that $r' \in C_{s'}$ and substitute the first row of

$$\left(\times \; (-B_{r'j})_{j \in C_1} \; \cdots \; (\mathbf{I} \cdot \delta_{r'j} - B_{r'j})_{j \in C_{s'}} \; \cdots (-B_{r'j})_{j \in C_m}\right)$$

by the sum of the rows $i \in B_{r'}$. Since for every s'' it holds that $Q_{r',a_{j_1}} = Q_{r',a_{j_2}}$ for all $j_1, j_2 \in \bar{C}_{s''}$ this sum has the form

$$\left(\times \; (-xQ_{r',a_j})_{j \in \bar{C}_1} \; \cdots \; (\bar{\delta}_{r'j} - xQ_{r',a_j})_{j \in \bar{C}_{s'}} \; \cdots (-xQ_{r',a_j})_{j \in \bar{C}_m}\right),$$

where $\bar{\delta}_{r'j} = 1$ if and only if $j \in B_{r'}$ and $= 0$ otherwise. This means that for every $s'' \neq s'$ the entries $(-xQ_{r',a_j})_{j \in \bar{C}_{s''}}$ are either all equal or, if $s'' = s'$, then we have to add 1 at proper positions. For every s'' we now add row $\mathbf{w}_{s''}$ xQ_{r',a_j} times. If $s'' \neq s'$ then we get a zero block $(0, \ldots, 0)$. If $s'' = s'$ we get a block of the form

$$\left(0 \cdots 0 \; \cdots \; 1 \cdots 1 \; \cdots \; 0 \cdots 0\right).$$

This means that this row is replaced by

$$\mathbf{w}_{s',r'} = \left(\times \; | \; 0 \cdots 0 \; | \; \cdots \; | \; 0 \cdots 0 \; | \; 0 \cdots 0 \; \cdots \; 1 \cdots 1 \; \cdots \; 0 \cdots 0 \; | \; 0 \cdots 0 \; | \; \cdots \; | \; 0 \cdots 0\right).$$

With the help of these rows we can eliminate all further entries of $\mathbf{M}(x, \mathbf{a})$ that come from $\mathbf{G}(x, \mathbf{a})$. (Here we use the fact that each row of $B_{r'r''}$ is either zero or all entries are the same.) This means that we finally end up with a matrix of the form

$$\mathbf{H} = \begin{pmatrix} 1 & 1 \cdots & \cdots & \cdots 1 \\ \times & H_{11} & \cdots & H_{1m} \\ \vdots & \vdots & & \vdots \\ \times & H_{m1} & \cdots & H_{mm} \end{pmatrix},$$

where $H_{s's''} = \mathbf{0}$ for $s' \neq s''$ and $H_{s's'}$ is of the form

$$H_{s's'} = \begin{pmatrix} J & K & K & \cdots & K \\ \mathbf{0} & J & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & J \end{pmatrix}.$$

with

$$J = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

It is now an easy task to transform the matrix $(H_{s's''})_{1\leqslant s',s''\leqslant m}$ (with the help of row transforms) to the identity matrix. Furthermore, we can transform the very first row $(1,1,\dots,1)$ of $\mathbf{H}$ to $(1,0,\dots,0)$ and end up with a matrix of the form

$$\begin{pmatrix} 1 & 0 \cdots & 0 \\ \times & 1 & 0 \\ \vdots & \ddots & \vdots \\ \times & 0 & 1 \end{pmatrix}.$$

Obviously, this matrix has determinant 1. Since the above row transforms do not change the value of the determinant, we thus obtain $\det \mathbf{M}(x,\mathbf{a}) = 1$.

Acknowledgement

The authors want to thank Philippe Flajolet for several discussions on the topic of the paper and for many useful hints.

References

[1] Aho, A. V., Sethi, R. and Ullman, J. D. (1986) *Compilers: Principles, Techniques, and Tools*, Addison-Wesley.

[2] Bender, E. A. and Richmond, L. B. (1983) Central and local limit theorems applied to asymptotic enumeration II: Multivariate generating functions. *J. Combin. Theory Ser. A* **34** 255–265.

[3] Drmota, M. (1994) Asymptotic distributions and a multivariate Darboux method in enumeration problems. *J. Combin. Theory Ser. A* **67** 169–184.

[4] Drmota, M. (1997) Systems of functional equations. *Random Struct. Alg.* **10** 103–124.

[5] Drmota, M. and Gittenberger, B. (1999) The distribution of nodes of given degree in random trees. *J. Graph Theory* **31** 227–253.

[6] Dershowitz, N. and Zaks, S. (1989) Patterns in trees. *Discrete Appl. Math.* **25** 241–255.

[7] Flajolet, P., Gourdon, X. and Martínez, C. (1997) Patterns in random binary search trees. *Random Struct. Alg.* **11** 223–244.

[8] Flajolet, P. and Sedgewick, R. (2006) *Analytic Combinatorics.* `http://algo.inria.fr/flajolet/Publications/books.html`

[9] Flajolet, P. and Steyaert, J.-M. (1980) On the analysis of tree-matching algorithms. In *Automata, Languages and Programming: Proc. Seventh Internat. Colloq., Noordwijkerhout, 1980* (G. P. Bhattacharjee, ed.), Vol. 85 of *Lecture Notes in Computer Science*, Springer, Berlin, pp. 208–219.

[10] Flajolet, P. and Steyaert, J.-M. (1980) On the analysis of tree-matching algorithms. In *Trees in Algebra and Programming: Proc. 5th Lille Colloq., Lille, 1980*, University of Lille I, pp. 22–40.

[11] Kok, G. (2005) Pattern distribution in various types of random trees. In *2005 International Conference on Analysis of Algorithms*, pp. 223–230.

[12] Kok, G. J. P. (2005) The distribution of patterns in random trees. Thesis, Institut für Diskrete Mathematik und Geometrie, TU Wien, Austria.
[13] Lalley, S. Random walks on infinite free products and infinite algebraic systems of generating functions. http://www.stat.uchicago.edu/~lalley/Papers.
[14] Meir, A. and Moon, J. W. (1978) On the altitude of nodes in random trees. *Canad. J. Math.* **30** 997–1015.
[15] Otter, R. (1948) The number of trees. *Ann. of Math.* (2) **49** 583–599.
[16] Robinson, R. W. and Schwenk, A. J. (1975) The distribution of degrees in a large random tree. *Discrete Math.* **12** 359–372.
[17] Ruciński, A. (1988) When are small subgraphs of a random graph normally distributed? *Probab. Theory Rel. Fields* **78** 1–10.
[18] Steyaert, J.-M. and Flajolet, P. (1983) Patterns and pattern-matching in trees: An analysis. *Inform. Control* **58** 19–58.