Counting Embeddings of Rooted Trees into Families of Rooted Trees *

Bernhard Gittenberger[1], Zbigniew Gołębiewski[2], Isabella Larcher[1], and Małgorzata Sulkowska[2,3]

[1] Department of Discrete Mathematics and Geometry
Technische Universität Wien
Wien, Austria

gittenberger@dmg.tuwien.ac.at
isabella.larcher@gmail.com

[2] Department of Fundamentals of Computer Science
Wrocław University of Science and Technology
Wrocław, Poland

{zbigniew.golebiewski, malgorzata.sulkowska}@pwr.edu.pl

[3] Université Côte d'Azur, CNRS, Inria, I3S
Sophia-Antipolis, France

Submitted: Sep 29, 2021; Accepted: Aug 5, 2022; Published: Sep 9, 2022

Abstract

An embedding of a rooted tree S into another rooted tree T is a mapping of the vertex set of S into the vertex set of T that is order-preserving in a certain sense, depending on the chosen tree family. Equivalently, S and T may be interpreted as tree-like partially ordered sets, where S is isomorphic to a subposets of T. A good embedding is an embedding where the root of S is mapped to the root of T.

We investigate the number of good and the number of all embeddings of a rooted tree S into the family of all trees of given size n of a certain family of trees. The tree families considered are binary plane trees (the order of descendants matters), binary non-plane trees and rooted plane trees. We derive the asymptotic behaviour of the number of good and the number of all embeddings in all these cases and prove that the ratio of the number of good embeddings to that of all embeddings is of the order $\Theta(1/\sqrt{n})$ in all cases, where we provide the exact constants as well. Furthermore,

*This research has been supported by the ÖAD, grant PL04-2018, and Wrocław University of Science and Technology grant 0401/0052/18.

we prove some monotonicity properties of this ratio. Finally, we comment on the case where S is disconnected.

Mathematics Subject Classifications: 05C05, 05A15, 05A16, 60G40, 06A07

1 Introduction

This paper studies the number of occurrences of a given rooted tree in the family of (plane and non-plane) binary trees, as well as planted plane trees. We call a rooted tree *plane* if the descendants of any vertex are ordered, and *non-plane* otherwise. Here, the notion of occurrence is wider than just a copy. We call it an *embedding* and we distinguish between plane embeddings (addressed to plane structures) and non-plane embeddings (for non-plane structures). In the non-plane case it is convenient to explain the notion of embedding we work with in terms of partially ordered sets (in short: *posets*). Recall that a partial order is a reflexive, antisymmetric and transitive relation. Note that any rooted tree may be interpreted as a Hasse diagram of some poset. Thus in the non-plane case we assume the investigated structures to be posets and by saying that there exists an embedding of S into T we understand that a poset S is a subposet of T. For the plane case assume that S and T are both rooted trees. Then a plane embedding of S into T is any subposet of T isomorphic to S in which the order of descendants of each node of S is inherited from T. We also distinguish between good embeddings in which the roots of S and T coincide and bad embeddings in which they do not.

The number of good and bad embeddings of a rooted structure in a complete binary tree was first investigated by Morayne [37]. His research was motivated by optimal stopping problems. The ratio of the number of good embeddings to the number of all embeddings and its monotonicity properties were used in estimates of conditional probabilities needed to obtain an optimal policy for the *best choice problem* considered on a complete (balanced) binary tree. This and similar results first served just as tools but soon became interesting questions about the structural features of posets on their own and resulted in a series of self-standing papers [29, 30, 22]. Counting chains and antichains in trees took a special place in this pool [33, 34, 31].

In this paper we present a follow-up and generalization of the results obtained by Kubicki *et al.* [29, 30] and Georgiou [22]. We give the asymptotic behaviour of the number of good and all embeddings of a rooted tree S in the family of plane and non-plane binary trees, as well as planted plane trees, on n vertices. The ratio of the number of good embeddings to the number of all embeddings is shown to be of order $\Theta(1/\sqrt{n})$ in all cases and the exact constants are provided. Furthermore, we show that this ratio is asymptotically non-decreasing in S (that is, if S_1 may be embedded in S_2, then the asymptotic ratio of good to all embeddings of S_1 is smaller or equal to the one for S_2). Such a monotonicity property was the center of attention of the aforestated papers. The detailed discussion on this matter is presented in Section 8. To obtain our results we use an approach based upon analytic combinatorics, which has not been used so far in the realm of posets.

The results of our paper may also be put into the framework of counting patterns in large structures. This is a vast field where many different types of structures have been considered. We only mention subgraph avoidance (and characterizing whole graph classes like series-parallel or planar graphs in that way) or subgraph counts in random graphs (see [27, 1] for general graphs, [12] for subcritical graphs and [38, 46] for planar graphs), pattern avoidance in permutations (see [5]) or in trees (see [10]), or pattern avoidance in lattice paths and words, where many particular patterns have been treated separately (see [9] or the introduction of [2] for a survey) and eventually put under a unifying umbrella in [2].

The closest to the present work is pattern counting in trees. One of the earliest investigations of this kind was [42], where the enumeration of given stars as subgraphs in trees (equivalently nodes of fixed degree) was treated. Later generalizations are found in [11, 40] (multivariate setting), in [35] (distinct patterns) or [24] (large patterns of that type). A method to deal with general contiguous patterns in trees by means of generating functions was developed in [6], which was partially generalized to planar maps recently [14, 7, 13]. Pattern avoidance in trees was the topic of [43], where also the concept of Wilf equivalence was dealt with, which was adopted from pattern avoidance in permutations.

Except for permutations, where most of the patterns that have been studied so far are non-contiguous, the considered patterns in other domains are typically contiguous. To our knowledge, the first work considering non-contiguous patterns in trees is [8]. In the present paper, the tree which is embedded becomes in general a collection of (partially) non-adjacent nodes in the tree where it is embedded. It can therefore be seen as a non-contiguous pattern occurring in that tree. Thus, our paper deals with certain enumeration problems for non-contiguous patterns in trees.

The paper is organized as follows. Section 2 introduces basic definitions, notation and presents tools (mainly from analytic combinatorics) used throughout the paper. Section 3 provides possible applications of our results in optimal stopping problems. In Section 4, using the symbolic method, we obtain generating functions for the number of good and the number of all embeddings of a rooted tree in the family of all plane binary trees with a given number of vertices. We also briefly discuss the case when the embedded structure is disconnected, *i.e.* it is a forest. Next, in Section 5, we use singularity analysis to derive the asymptotics of the number of good and bad embeddings when the size of the underlying tree is tending to infinity. Sections 6 and 7 present analogous results for the families of non-plane binary trees and planted plane trees, respectively. Moreover, in Section 8 we investigate the asymptotics and the monotonicity of the ratio of good to all embeddings of a rooted tree in all three mentioned families, which is a continuation of work of Kubicki *et al.* [29, 30] and Georgiou [22]. A discussion of the obtained results as well as an outlook into some related future problems is given in Section 9.

2 Preliminaries, methodology and concepts

We are dealing with several tree structures that will be presented in the following subsection. In all classes that we are studying the trees are considered as unlabelled graphs.

The subsequent subsections provide the methodology from analytic combinatorics. Here we present only a brief account of the symbolic method and singularity analysis. For a detailed presentation of the theory see [16].

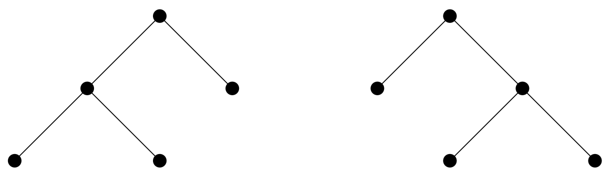
Finally, we present the definitions of all the concepts that are central to our investigations.

2.1 Structures

By $\mathcal{B}_n$ we denote the family of *plane binary trees* with n nodes. A binary tree is a tree in which each node has either 0 or 2 descendants and by plane we understand that the order of subtrees of a given node matters, *i.e.* we distinguish between the different embeddings of a tree in the plane. It is commonly known that for odd n the cardinality $|\mathcal{B}_n|$ of $\mathcal{B}_n$ satisfies $|\mathcal{B}_n| = \boldsymbol{C}_{\frac{n-1}{2}}$, where $\boldsymbol{C}_k$ is the k-th Catalan number given by $\boldsymbol{C}_k = \frac{1}{k+1}\binom{2k}{k}$. Note that all binary trees have odd sizes and thus, for even n the cardinality $|\mathcal{B}_n|$ is zero. All plane binary trees of size 5 are shown in Figure 1. We assume also that all edges are directed towards the descendants. Therefore, the in-degree of the root, as well as the out-degree of each leaf, is always 0. A vertex is said to be d-ary if its out-degree equals d. Subsequently, the root of a tree T will be denoted by $\mathbb{1}_T$.

By $\mathcal{V}_n$ we denote the family of *non-plane binary trees* with n nodes. By non-plane we understand that the subtrees of a given node are treated as a set of subtrees, *i.e.* there is no ordering. E.g., there is only one non-plane binary tree of size 5, see Figure 1. Again for even n the cardinality $|\mathcal{V}_n|$ is zero. For odd n the values $|\mathcal{V}_n|$ are known as Wedderburn-Etherington numbers and do not have a closed form ($|\mathcal{V}_1| = 1, |\mathcal{V}_3| = 1, |\mathcal{V}_5| = 1, |\mathcal{V}_7| = 2, |\mathcal{V}_9| = 3, \ldots$).

Planted plane trees (also known as Catalan trees) are rooted plane trees where each internal node can have arbitrarily many descendants. We denote the family of planted plane trees of size n by $\mathcal{T}_n$. For all n the cardinality of $\mathcal{T}_n$ satisfies $|\mathcal{T}_n| = \boldsymbol{C}_{n-1}$.



T_1 T_2

Figure 1: The family $\mathcal{B}_5 = \{T_1, T_2\}$ of plane binary trees is of size $|\mathcal{B}_5| = \boldsymbol{C}_2 = 2$, while the family $\mathcal{V}_5 = \{T_1\}$ of non-plane binary trees has the size $|\mathcal{V}_5| = 1$.

2.2 Tools from analytic combinatorics – the symbolic method

Let $\mathbb{N}$ denote the set of natural numbers including 0. A *combinatorial class* $\mathcal{A}$ is a pair $(A, |\cdot|_{\mathcal{A}})$ where A is a countable set whose elements are called (combinatorial) objects and $|\cdot|_{\mathcal{A}} : A \to \mathbb{N}$ is the *size function* of $\mathcal{A}$. We require that for all $n \in \mathbb{N}$ the set of all objects of size n has finite cardinality a_n. The *counting sequence* of $\mathcal{A}$ is the sequence $(a_n)_{n\geqslant 0}$.

Two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$ are said to be *combinatorially isomorphic* (or, simply, *isomorphic*) if and only if their counting sequences are identical. Then we write $\mathcal{A} \cong \mathcal{B}$. The *generating function* of a combinatorial class $\mathcal{A}$ is the generating function of its counting sequence $(a_n)_{n\geqslant 0}$, *i.e.* the formal power series $A(z) = \sum_{n\geqslant 0} a_n z^n$. By $[z^n]A(z)$ we denote the coefficient of z^n in the formal power series $A(z)$; thus $[z^n]A(z) = a_n$.

The *neutral class* $\mathcal{E}$ is defined as the class consisting of a single object of size 0; the *atomic class* containing only a single object of size 1 will be denoted by $\{\bullet\}$. The following combinatorial constructions for combinatorial classes are going to be used throughout the paper. We list them together with the relations between the corresponding generating functions.

1. **Sum (disjoint union).** Let $\mathcal{B} = (B, |\cdot|_{\mathcal{B}})$ and $\mathcal{C} = (C, |\cdot|_{\mathcal{C}})$ be combinatorial classes such that $B \cap C = \varnothing$. Then $\mathcal{A} \cong \mathcal{B} + \mathcal{C}$ if and only if $\mathcal{A} = (A, |\cdot|_{\mathcal{A}})$ where $A = B \cup C$ and $|\omega|_{\mathcal{A}} = |\omega|_{\mathcal{B}}$ if $\omega \in B$ and $|\omega|_{\mathcal{A}} = |\omega|_{\mathcal{C}}$ if $\omega \in C$. If $\mathcal{A} \cong \mathcal{B} + \mathcal{C}$ then $A(z) = B(z) + C(z)$.

2. **Product.** Let $\mathcal{B} = (B, |\cdot|_{\mathcal{B}})$ and $\mathcal{C} = (C, |\cdot|_{\mathcal{C}})$ be combinatorial classes. Then $\mathcal{A} \cong \mathcal{B} \times \mathcal{C}$ if and only if $\mathcal{A} = (A, |\cdot|_{\mathcal{A}})$ where $A = B \times C$ (the Cartesian product) and $|(b, c)|_{\mathcal{A}} = |b|_{\mathcal{B}} + |c|_{\mathcal{C}}$. If $\mathcal{A} \cong \mathcal{B} \times \mathcal{C}$ then $A(z) = B(z) \cdot C(z)$, where $\cdot$ denotes the Cauchy product.

3. **Sequence.** Let $\mathcal{A}$ be a combinatorial class with no object of size 0. We define
$$\mathcal{B} := \mathcal{S}eq(\mathcal{A}) = \mathcal{E} + \mathcal{A} + (\mathcal{A} \times \mathcal{A}) + (\mathcal{A} \times \mathcal{A} \times \mathcal{A}) + \cdots .$$
Then $B(z) = 1/(1 - A(z))$.

4. **Multiset of size 2.** Let $\mathcal{A}$ be a combinatorial class. Then we define
$$\mathcal{B} := \mathcal{MS}et_2(\mathcal{A}) = (\mathcal{A} \times \mathcal{A})/\mathbf{R},$$
with $\mathbf{R}$ being the equivalence relation defined by $(\alpha_1, \alpha_2)\mathbf{R}(\beta_1, \beta_2)$ if and only if $(\beta_1, \beta_2) = (\alpha_1, \alpha_2)$ or $(\beta_1, \beta_2) = (\alpha_2, \alpha_1)$. The multiset $\{a, b\}$, $a \neq b$, is associated with the two ordered pairs (a, b) and (b, a) whereas $\{a, a\}$ is associated with (a, a). The set of all pairs (a, a) is called the diagonal of $\mathcal{A} \times \mathcal{A}$ and denoted by $\Delta(\mathcal{A} \times \mathcal{A})$. The above explained correspondence translates to the following combinatorial isomorphism,
$$\mathcal{MS}et_2(\mathcal{A}) + \mathcal{MS}et_2(\mathcal{A}) \cong \mathcal{A} \times \mathcal{A} + \Delta(\mathcal{A} \times \mathcal{A}).$$
As the generating function of $\Delta(\mathcal{A} \times \mathcal{A})$ is $A(z^2)$, we get
$$B(z) = \frac{1}{2}\left(A(z)^2 + A(z^2)\right).$$
We remark that a similar argument works for sets. With $\mathcal{C} = \mathcal{S}et_2(\mathcal{A})$ we have
$$C(z) = \frac{1}{2}\left(A(z)^2 - A(z^2)\right). \tag{1}$$

5. **Pointing.** Let $\mathcal{A}$ be a combinatorial class. Pointing of a class $\mathcal{A}$ means distinguishing an atom (atoms are objects of size 1) and objects with a distinguished atom are called pointed objects. Precisely, the pointed class is the class of all pointed objects made from an object of $\mathcal{A}$. It is denoted by $\mathcal{B} = \Theta(\mathcal{A})$ and formally defined as

$$\Theta(\mathcal{A}) = \sum_{n\geqslant 0} \mathcal{A}_n \times \{\epsilon_1, \ldots, \epsilon_n\},$$

where $\{\epsilon_1, \ldots, \epsilon_n\}$ is a fixed collection of distinct neutral objects of size 0. With a_n being the number of objects of size n in $\mathcal{A}$, the quantity na_n is apparently the number of pointed objects of size n in $\mathcal{B}$, as each of the n atoms may be pointed at. In terms of generating functions the relation reads as $B(z) = zA'(z)$.

2.3 Trees

From now on let $\mathcal{B}$ denote the class of plane binary trees, $\mathcal{V}$ the class of non-plane binary trees, $\mathcal{M}$ the class of Motzkin trees (defined below) and $\mathcal{T}$ the class of planted plane trees. Let also $B(z) = \sum_{n\geqslant 0} b_n z^n$, $V(z) = \sum_{n\geqslant 0} v_n z^n$, $M(z) = \sum_{n\geqslant 0} m_n z^n$ and $T(z) = \sum_{n\geqslant 0} t_n z^n$ be their corresponding generating functions where z marks the number of nodes (e.g., b_n is the number of plane binary trees with n nodes), thus

$$\begin{aligned} B(z) &= \boldsymbol{C}_0 z + \boldsymbol{C}_1 z^3 + \boldsymbol{C}_2 z^5 + \boldsymbol{C}_3 z^7 + \cdots\ , \\ V(z) &= z + z^3 + z^5 + 2z^7 + 3z^9 + \cdots\ , \\ M(z) &= z + z^2 + 2z^3 + 4z^4 + 9z^5 + \cdots\ , \\ T(z) &= \boldsymbol{C}_0 z + \boldsymbol{C}_1 z^2 + \boldsymbol{C}_2 z^3 + \cdots\ . \end{aligned}$$

Each of the classes may be specified by a recursive relation using the symbolic method. Obviously, we have

$$\mathcal{B} = \{\bullet\} + \{\bullet\} \times \mathcal{B} \times \mathcal{B},$$

since a plane binary tree consists either of a single vertex or of a root with two plane binary trees attached to it. Translating into generating functions gives the quadratic equation $B(z) = z + zB(z)^2$ with one of its solutions being a power series. Thus

$$B(z) = \frac{1 - \sqrt{1 - 4z^2}}{2z}.$$

A Motzkin tree is a plane rooted tree in which each vertex has either zero, one or two children that are themselves Motzkin trees. This yields the specification

$$\mathcal{M} = \{\bullet\} + \{\bullet\} \times \mathcal{M} + \{\bullet\} \times \mathcal{M} \times \mathcal{M}. \tag{2}$$

Similarly, we define planted plane trees by

$$\mathcal{T} = \{\bullet\} \times \mathcal{S}eq(\mathcal{T}).$$

The specifications of Motzkin trees and of planted plane trees lead to the quadratic equations $M(z) = z + zM(z) + zM(z)^2$ and $T(z) = z/(1 - T(z))$, respectively, giving rise to

$$M(z) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z} \quad \text{and} \quad T(z) = \frac{1 - \sqrt{1 - 4z}}{2}.$$

In non-plane binary trees, the two subtrees attached to the root do not form a pair, but a multiset instead. This yields

$$\mathcal{V} = \{\bullet\} + \{\bullet\} \times \mathcal{MS}et_2(\mathcal{V})$$

and so

$$V(z) = z + \frac{z}{2}(V(z)^2 + V(z^2)). \tag{3}$$

2.4 Singularity analysis and other auxiliary results

We will use singularity analysis to get a relation between the behaviour of a generating function near its dominant singularities (*i.e.* its singularities on the circle of convergence) and the asymptotics of its coefficients. Precisely, we take advantage of the following lemma. (We use the standard notation $f(n) \sim g(n)$ if $\lim_{n \to n_0} \frac{f(n)}{g(n)} = 1$.)

Lemma 1 (Compare Theorems VI.4 and VI.5 in [16]). *Define*

$$\Delta_0 = \{z \in \mathbb{C} | |z| < \rho + \epsilon, z \neq \rho, |\mathrm{arg(z - \rho)}| > \nu\}$$

for some $\rho > 0, \epsilon > 0, 0 < \nu < \frac{\pi}{2}$. Let $r \geqslant 0$, $\rho_j = \rho e^{i\phi_j}$, for $j = 0, 1, \ldots, r$ with $\phi_0 = 0$ and $\phi_1, \ldots, \phi_r \in (0, 2\pi)$. Consider $T(z) = \sum_{n \geqslant 0} T_n z^n$ to be an analytic function in $\Delta := \bigcap_{j=0}^{r} e^{i\phi_j} \Delta_0$ and satisfying for each $j = 0, \ldots, r$

$$T(z) \sim K_j \left(1 - \frac{z}{\rho_j}\right)^{-\alpha_j}, \qquad \text{as } z \to \rho_j \text{ in } \Delta,$$

where $\alpha_j \notin \{0, -1, -2, \ldots\}$ and the K_j are constants. Then

$$[z^n]T(z) \sim \sum_{j=0}^{r} K_j \frac{n^{\alpha_j - 1}}{\Gamma(\alpha_j)} \rho_j^{-n}, \qquad \text{as } n \to \infty.$$

Remark 2. Note that the assumptions of Lemma 1 imply that $\{\rho_0, \rho_1, \ldots, \rho_r\}$ is exactly the set of all singularities of the power series $\sum_{n \geqslant 0} T_n z^n$ on its circle of convergence.

Finally, the following lemma will be helpful when investigating the asymptotic monotonicity of ratios of good to all embeddings.

Lemma 3 (Gautschi's inequality, [21]). *Let x be a positive real number and let $s \in (0, 1)$. Then*

$$x^{1-s} < \frac{\Gamma(x + 1)}{\Gamma(x + s)} < (x + 1)^{1-s}.$$

2.5 Embeddings

This paper concentrates on investigating the number of embeddings of any rooted tree (or a forest of rooted trees - a disconnected graph whose components are rooted trees) in all trees from either family $\mathcal{B}_n$, $\mathcal{V}_n$ or $\mathcal{T}_n$. An embedding of a rooted tree S into another rooted tree T can be seen as a kind of generalized pattern occurrence of S in T, where we distinguish between the plane and the non-plane case. As we are extending some results on posets, we interpret S and T as posets and formulate the definitions in the language of posets (see e.g. [48] for basic concepts and terminology), but give the equivalent definition in terms of trees afterwards.

Definition 4 (*Non-plane embedding)***.** Let S and T be two non-plane rooted trees. When interpreting T as the cover graph of a poset, rooted at the root of T, *i.e.* at the single maximal element of the poset, then an embedding of S into T can be defined is any subposet of T that is isomorphic to S.

Definition 4' (*Non-plane embedding)***.** Let S and T be two non-plane rooted trees with vertex sets $V(S)$ and $V(T)$, respectively. Then an embedding of S into T is a subset $M \subseteq V(T)$ for which there is a bijection $\varphi : M \to V(S)$ such that x is a descendant of y in T if and only if $\varphi(x)$ is a descendant of $\varphi(y)$ in S.

Remark 5. Note that there exists a non-plane embedding of a binary tree S into a binary tree T if and only if S is a minor of T.

Remark 6. Instead of starting from a tree as combinatorial structure and then interpreting it as a poset, we may also start from posets and then define a *tree poset* as a poset P which has exactly one maximal element and such that any Hasse diagram of P looks like a (combinatorial) tree. This is equivalent to the definition of a tree poset given in [19].[1] Likewise, an embedding of a tree poset S into another tree poset T, as defined in [19], matches exactly the definition of a non-plane embedding given above.

Definition 7 (*Plane embedding)***.** Let S and T be two plane rooted trees. If we interpret T to be a Hasse diagram of a poset, then an embedding of S into T is defined as any subposet of T isomorphic to S in which the left-to-right order of the children of each node of S is inherited from T (one may think of it as of a plane version of a subposet).

Remark 8. So, in the plane case S and T can be interpreted as Hasse diagrams of posets, and whenever S can be embedded in T it follows that S is a subposet of T. However, note that the respective posets can possibly be represented as different Hasse diagrams in such a way that no embedding of the corresponding trees is possible.

Definition 7' (*Plane embedding)***.** Let S and T be two plane rooted trees with vertex sets $V(S)$ and $V(T)$, respectively. Then an embedding of S into T is a subset $M \subseteq V(T)$ for which there is a bijection $\varphi : M \to V(S)$ meeting the following two constraints: x is

[1]For the sake of better distinction from a combinatorial tree, we use the term "tree poset" for what is simply called "tree" in [19].

a descendant of y in T if and only if $\varphi(x)$ is a descendant of $\varphi(y)$ in S; if for two vertices v and w the left subtree of their last common ancestor z contains v (and thus w is in the right subtree attachted to z), then $\varphi(v)$ is in the left subtree of the last common ancestor u of $\varphi(v)$ and $\varphi(w)$.

Note here: Usually $\varphi^{-1}(u) \neq z$, yet in general we even have $z \notin M$.

We say that an embedding of S into T is *good* if it contains the root of T. Otherwise we call it a *bad embedding*. If there exists at least one embedding of S into T, we write $S \subseteq T$. All embeddings of a cherry, (*i.e.* a tree composed only of a root and its two children) in a given binary tree of size 5 are given in Figure 2. Four of them are good and the last one is bad.

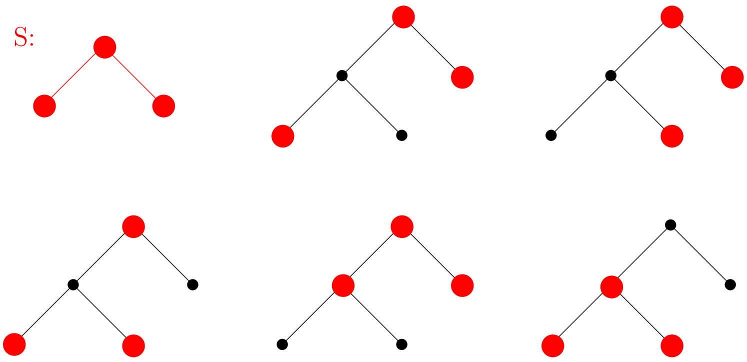


Figure 2: All five plane embeddings of a cherry S in a given plane binary tree of size 5. Or all four non-plane embeddings of a cherry S in a given non-plane binary tree of size 5, since in the non-plane case the two rightmost pictures in the upper row represent the same embedding (they can easily be mapped onto each other via a simple automorphism that changes the order of the two leftmost leaves).

Subsequently, the size of the tree S will always be denoted by m, while the size of T is consistently denoted by n. Thus, for the asymptotic analysis of the number of embeddings of a tree S into a class of trees of size n, the quantity m is considered to be a constant, while n tends to infinity.

For S, the structure that we embed, we define its *degree distribution sequence* as $d_S = (d_0, d_1, \dots, d_{m-1})$, where d_i is the number of vertices in S with out-degree equal to i. Note that d_0 is simply the number of leaves, which will be, interchangeably, denoted by l (*i.e.* $l = d_0$). Similarly, d_1 is the number of unary nodes, which will be, interchangeably, denoted by u (*i.e.* $u = d_1$). The number of all embeddings of a given tree S in T will be denoted by $a_T(S)$ and the number of its good embeddings in T by $g_T(S)$.

The number of all embeddings of S in a finite family $\mathcal{F} = \{F_1, \dots, F_\ell\}$ will be denoted by $a_{\mathcal{F}}(S)$ and understood as the cumulative number of embeddings of S into all elements of $\mathcal{F}$, *i.e.* $a_{\mathcal{F}}(S) = \sum_{i=1}^{\ell} a_{F_i}(S)$. Analogously, we define the number of good embeddings of S in $\mathcal{F}$: $g_{\mathcal{F}}(S) = \sum_{i=1}^{\ell} g_{F_i}(S)$. For S being a cherry and $\mathcal{F} = \mathcal{B}_5 = \{T_1, T_2\}$ from Figure 1, we obtain $a_{T_1}(S) = a_{T_2}(S) = 5$, $g_{T_1}(S) = g_{T_2}(S) = 4$, thus $a_{\mathcal{B}_5}(S) = 10$ and $g_{\mathcal{B}_5}(S) = 8$ (compare Figure 2).

Moreover, for each considered family $\mathcal{F}$ of trees with $\mathcal{F}_n$ denoting the subfamily of all trees of size n, we define the generating functions of the numbers of all and good embeddings of a structure S as

$$A_S(z) = \sum_{n \geqslant 0} a_{\mathcal{F}_n}(S) z^n \qquad \text{and} \qquad G_S(z) = \sum_{n \geqslant 0} g_{\mathcal{F}_n}(S) z^n.$$

3 Applications in optimal stopping problems

The most prominent problem in the area of optimal stopping is the so-called "secretary problem" (consult [36, 15, 17, 44]), where one assumes a linear order on the applicants for a secretary position concerning their qualifications. The applicants are interviewed in a random order and the decision whether to hire an applicant has to be made immediately after the interview - a rejected applicant cannot be hired at a later point. Thus, if we interview all the candidates, we have to hire the last applicant. The goal is to find the optimal stopping strategy to hire the best applicant. Thus, we want to stop at the time maximizing the probability that the present applicant is the best one overall, *i.e.* the maximum element in the linear order. It has been proved (see for example [36, 23]) that for a large number of applicants it is optimal to wait until approximately 37% (more precisely $\frac{100}{e}$%) of the applicants have been interviewed and then to select the next relatively best one. This optimal algorithm returns the best applicant with asymptotic probability of $1/e$. The secretary problem has been extended and generalized in many different directions. One of these is the extension to partially ordered sets, possibly with more than one maximal element, see [45, 25]. Optimal strategies for particular posets were investigated among others in [37, 28]. Versions for unknown poset, when the selector knows in advance only its cardinality, were presented in [41, 18, 20]. Another interesting generalization was to replace the underlying poset structure by a directed graph. This version was first considered on directed paths by Kubicki and Morayne in [32] and later extended to other families of graphs and different versions of the game (consult [47, 26, 3, 4]).

In the remainder of this section we give examples of stopping problems in which either the value $a_{\mathcal{V}_n}(S)$ or the ratio $g_{\mathcal{V}_n}(S)/a_{\mathcal{V}_n}(S)$ (both investigated in this paper) plays a crucial role in estimating the conditional probabilities needed to obtain the optimal policy. One can consider analogous examples for the families $\mathcal{B}_n$ or $\mathcal{T}_n$ as well.

Let us think about elements of $\mathcal{V}_n$ as of Hasse diagrams of posets. Consider the following process. Elements (*i.e.* nodes) of some T from $\mathcal{V}_n$ appear one by one in a random order (all permutations of elements of T are equiprobable). At time t, *i.e.* when t elements have already appeared, the selector can see a poset induced on those elements. He knows that the underlying structure is drawn uniformly at random from $\mathcal{V}_n$.

Example (Best choice problem for the family of binary trees)**.** The selector's task is to stop the process maximizing the probability that the element that has just appeared is the root of the underlying structure. He wins only if the chosen element is indeed $\mathbb{1}_T$. Note that it neither pays off to stop the process when the induced structure is disconnected nor when the currently observed element is not the maximal one in the induced poset.

The selector wonders whether to stop only if the emerged element at time t is the unique maximal element in the induced structure. In order to take a decision whether to stop at time t, he needs to know the probability of winning if he stops now. Let W_t denote the event of winning when stopping at time t, S_t the event that at time t he observes a certain structure S with degree distribution sequence d_S and R_i denote the event that T_i has been drawn as the underlying structure, where we use the notation $\mathcal{V}_n = \{T_1, \dots, T_N\}$ with $N = |\mathcal{V}_n|$. Then the probability of winning if he stops at time t is given by

$$\mathbb{P}[W_t|S_t] = \sum_{i=1}^{N} \mathbb{P}[W_t|S_t \cap R_i]\mathbb{P}[R_i|S_t] = \sum_{i=1}^{N} \frac{g_{T_i}(S)}{a_{T_i}(S)} \frac{\mathbb{P}[S_t|R_i]\mathbb{P}[R_i]}{\mathbb{P}[S_t]}.$$

Since $\mathbb{P}[R_i] = 1/N$, $\mathbb{P}[S_t|R_i] = a_{T_i}(S)/\binom{n}{t}$ (t vertices were revealed till time t; among all $\binom{n}{t}$ equiprobable choices of t vertices from T_i, $a_{T_i}(S)$ of them admit the embedding of S in T_i) and

$$\mathbb{P}[S_t] = \sum_{i=1}^{N} \mathbb{P}[S_t|R_i]\mathbb{P}[R_i] = \sum_{i=1}^{N} \frac{a_{T_i}(S)}{\binom{n}{t}} \frac{1}{N} = \frac{a_{\mathcal{V}_n}(S)}{N\binom{n}{t}}$$

we get

$$\mathbb{P}[W_t|S_t] = \sum_{i=1}^{N} \frac{g_{T_i}(S)}{a_{T_i}(S)} \frac{a_{T_i}(S)}{\binom{n}{t}} \frac{1}{N} \frac{N\binom{n}{t}}{a_{\mathcal{V}_n}(S)} = \frac{g_{\mathcal{V}_n}(S)}{a_{\mathcal{V}_n}(S)}.$$

Example (Identifying complete balanced binary trees)**.** The selector has to identify whether the underlying structure is a complete balanced binary tree or not. The payoff of the game, if he stops the process at time t, is $n-t$ if he guesses correctly and 0 otherwise. He has to maximize the expected payoff. At moment t he observes a structure S, which is not necessarily connected. Again, in order to make a decision whether to stop, he needs to know what is the probability that the currently observed structure is a subposet of a complete balanced binary tree. For a rooted tree S this probability is given by

$$\frac{a_{T_b}(S)}{a_{\mathcal{V}_n}(S)},$$

where $T_b \in \mathcal{V}_n$ denotes the complete balanced binary tree of size n.

4 Generating functions for the number of embeddings in $\mathcal{B}_n$

4.1 Embedding rooted trees in $\mathcal{B}_n$

In this subsection we derive explicit expressions for the generating functions for the sequences $a_{\mathcal{B}_n}(S)$ and $g_{\mathcal{B}_n}(S)$, where S is a given rooted plane tree of size m. In order to do so, we use the symbolic method as outlined in Section 2.

Theorem 9. *Consider a rooted tree S with $d_S = (l, u, d_2, \dots, d_{m-1})$ being its degree distribution sequence. The generating function $A_S(z)$ of the sequence $a_{\mathcal{B}_n}(S)$, which counts* *the number of all embeddings of S into all trees of the family $\mathcal{B}_n$, is given by*

$$A_S(z) = \left(\frac{1}{1-2zB(z)}\right)^{m+l-1} z^{l+u-1}\ B(z)^{l+u}\ 2^u \prod_{i=3}^{m-1} (\boldsymbol{C}_{i-1})^{d_i}, \tag{4}$$

where $B(z)$ is the generating function of the family of plane binary trees.

Remark 10. Note that $A_S(z)$ depends only on the degree distribution sequence d_S, not the particular shape of S. Thus, as long as d_{S_1} and d_{S_2} are the same, $A_{S_1}(z)$ and $A_{S_2}(z)$ coincide even if S_1 and S_2 are not isomorphic. However, we use the subscript S to provide a transparent notation. Moreover, note that $A_S(z)$ does also depend on the tree class $\mathcal{B}_n$ in which we embed the tree S. In order to avoid a large number of indices we will omit to indicate this dependence and just emphasize at this point that the generating functions $A_S(z)$ may differ according to the underlying tree classes.

Proof. **Case 1:** S is a Motzkin tree.

In the proof we heavily use the definitions class $\mathcal{B}$ of binary trees and the class $\mathcal{M}$ of Motzkin trees from Section 2. We start the proof of the expression of $A_S(z)$ with the case where S is a Motzkin tree. and thereby we must distinguish between the three cases whether S is a single node, or the root of S is a unary node, or a binary node, and hence falling into the respective subclass of $\mathcal{M}$ among the subclasses that we find as summands on the right-hand side of Eq. (2). In particular, if the root is a unary node, then there is one Motzkin tree attached to it that we call $\tilde{S}$. If the root is a binary node, it has a left and a right subtree attached to it that we call S_L and S_R, respectively.

The generating function $A_S(z)$ for the number of embeddings of S into the family $\mathcal{B}_n$ can then be recursively defined by

$$A_S(z) = \begin{cases} zB'(z) & \text{if } S = \bullet, \\ \dfrac{2zB(z)}{1-2zB(z)} A_{\tilde{S}}(z) & \text{if } S = (\bullet, \tilde{S}), \\ \dfrac{z}{(1-2zB(z))^2} A_{S_L}(z) A_{S_R}(z) & \text{if } S = (\bullet, S_L, S_R), \end{cases} \tag{5}$$

where the three cases correspond to the cases described above. The first case, which yields a factor $zB'(z)$, corresponds to marking a node in the underlying tree T (*i.e.* pointing at a node), because obviously a single vertex can be embedded in every node. We can also interpret it as counting the number of pairs (T, E) where E is an embedding of S into T.

Now we show how an embedding of S into T can be constructed in a recursive way - see Figure 3 for a visualization of the used approach. We start with the case that the root of S is a unary node. This root has to be embedded at some point in the tree T. The part of T that is above the embedded root of S can be expressed as a path of left-or-right trees, which contributes a factor $1/(1-2zB(z))$. The embedded root of S itself yields a factor z, since the generating function of an object of size one is given by z. To the embedded root we have to attach an additional tree T in order to create a binary structure, yielding a factor $B(z)$, as well as the remaining tree that contains the embedding of $\tilde{S}$. The factor

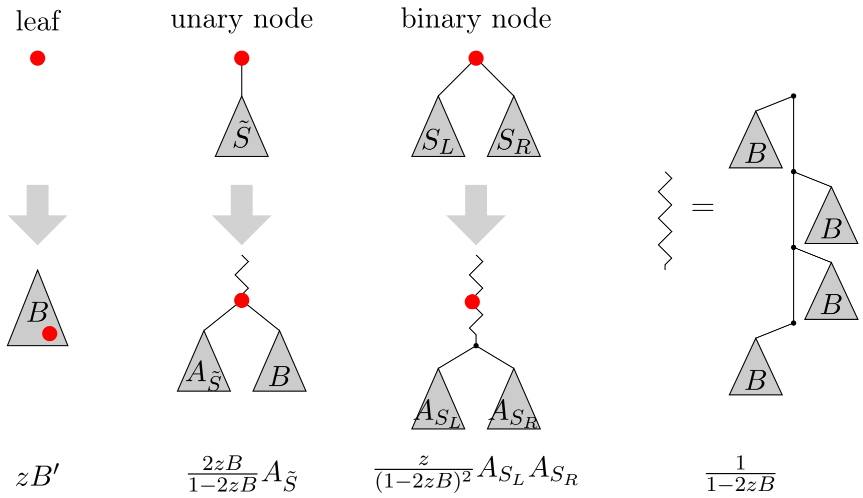


Figure 3: Sketch of the recursive construction of the generating function $A_S(z)$. If S is a Motzkin tree (plane binary case), then the three cases above can appear. Here B each time refers to an abstract object representing any tree from family $\mathcal{B}_n$.

2 that appears in the coefficient in the second case of (5) indicates that we work with plane trees - the substructure $\tilde{S}$ can be embedded either in the left or in the right subtree of the unary vertex.

The third case of (5), where S starts with a binary node, is similar to the previous case. Thus, the factor $1/(1-2zB(z))^2$ corresponds to two consecutive paths of left-or-right trees, which are separated by the embedded root which itself gives the additional factor z. At some point the lower path splits into two subtrees containing the embeddings of the subtrees S_L and S_R.

By simple iteration one can see that in case of embedding a Motzkin tree S, the generating function $A_S(z)$ reads as

$$A_S(z) = \left(\frac{z}{(1-2zB(z))^2}\right)^{l-1}\left(\frac{2zB(z)}{1-2zB(z)}\right)^{u}(zB'(z))^l, \tag{6}$$

where l denotes the number of leaves and u the number of unary nodes in S. The exponent $l-1$ in (6) arises from the fact that a Motzkin tree with l leaves has $l-1$ binary nodes, and for each of these nodes we get the respective factor.

Case 2: S contains vertices with more than two children.

Finally, we consider the general case where S is an arbitrary plane tree without any restrictions on the degree distribution sequence. Then we proceed as follows. Every d-ary node with $d \geqslant 3$ together with its d children is replaced by a binary tree having d leaves, which are then replaced by the successors of the original d-ary node. There are exactly $\boldsymbol{C}_{d-1}$ possible ways to construct such a binary tree. Unary and binary nodes stay unaltered. Applying this for all nodes results in constructing a Motzkin tree, called

S', and the number of Motzkin trees that can be constructed in that way is $\prod_{i=3}^{m-1} \boldsymbol{C}_{i-1}^{d_i}$. These Motzkin trees are then embedded with the approach described above.

In short, the goal is now to make S correspond to a collection of Motzkin trees $\mathcal{C} = \{S_1, \ldots, S_k\}$ with $k = \prod_{i=3}^{m-1} \boldsymbol{C}_{i-1}^{d_i}$ and each embedding of S into T could then be regarded as an embedding of some $S' \in \mathcal{C}$ and *vice versa*.[2] This would imply

$$A_S(z) = \sum_{i=1}^{k} A_{S_i}(z) = kA_{S'}(z),$$

where the second equation is true, because $A_{S_i}(z)$ only depends on the number of unary nodes and the number of leaves of S_i, but not on its actual shape.

In order to determine $A_{S'}$ note that replacing a d-ary node v with $d \geqslant 3$ by a binary tree (with more than 1 internal nodes) introduced further vertices into S. In particular, v becomes a binary node, and as there were d successors before, the binary tree arising from v must have $d-1$ internal nodes, thus giving rise to $d-2$ auxiliary vertices. Since S has m vertices, S' has therefore $m + \sum_{i=3}^{m-1}(i-2)d_i = 2l + u - 1$ vertices.

Furthermore note that the auxiliary vertices have to be embedded into T, but they do not belong to S. This causes a special treatment (see Figures 4 and 5 for an illustration). Consider two subtrees S_1 and S_2 of S' whose last common ancestor (in S') is an auxiliary vertex, say w, and let v be the last common ancestor of w and its sibling in S' (which exists, as auxiliary nodes only appear when dissolving vertices of arity 3 or more). Of course, any vertex on the path from v to w may serve as auxiliary vertex instead of w. But as w does not belong to S, its actual position is unimportant. Hence, for the sake of not overcounting, the auxiliary vertices are always placed at the last possible position, see Figure 5 for the local picture and Figure 4 for the global picture. This eventually yields a factor $z/(1-2zB(z))$ for each binary auxiliary node. As there are $2l+u-m-1$ auxiliary nodes and $m-l-u$ other binary nodes, this gives altogether

$$\begin{aligned} A_S(z) &= \left(\frac{z}{1-2zB(z)}\right)^{2l+u-m-1} \left(\frac{z}{(1-2zB(z))^2}\right)^{m-l-u} \\ &\quad \cdot \left(\frac{2zB(z)}{1-2zB(z)}\right)^{u} (zB'(z))^l \prod_{i=3}^{m-1} (\boldsymbol{C}_{i-1})^{d_i} \\ &= \frac{2^u z^{l+u-1} B(z)^u}{(1-2zB(z))^{m-1}} (zB'(z))^l \prod_{i=3}^{m-1} (\boldsymbol{C}_{i-1})^{d_i}. \end{aligned}$$

Using the identity $zB'(z) = B(z)/(1-2zB(z))$, which holds for plane binary trees, yields the desired result. □

[2] *Caveat*: S' has more vertices than S, and these must be embedded in a canonical, unique way. How this can be achieved is discussed in the subsequent paragraphs.

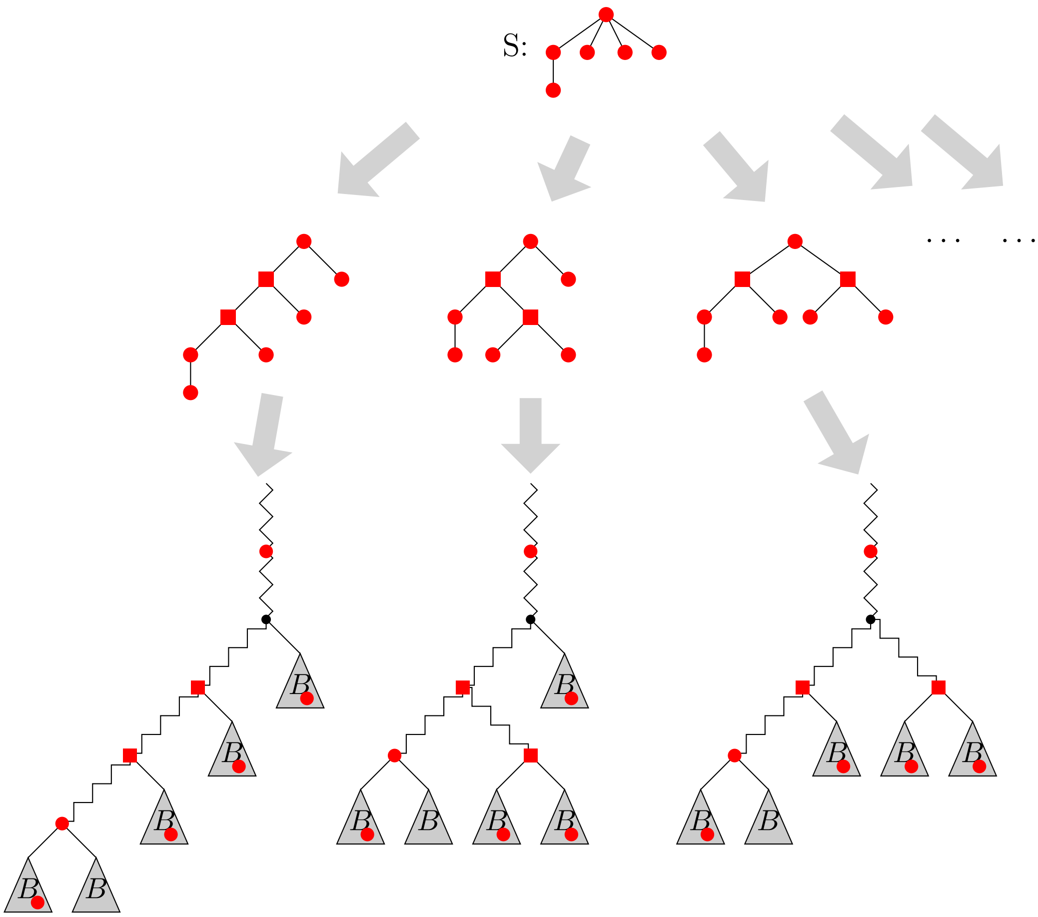


Figure 4: Sketch of the principle of embedding an arbitrary plane tree (plane binary case). When embedded into a binary structure, the quarternary node is replaced by one of the five possible plane binary trees, whereas the unary node and its child stay in their position, of course. The last two cases, not shown here, are symmetric to the first and second one.

Corollary 11. *Let S be a rooted tree. The generating function of the sequence $g_{\mathcal{B}_n}(S)$, which counts the number of good embeddings of S into all trees of the family $\mathcal{B}_n$, is given by*

$$G_S(z) = (1 - 2zB(z))A_S(z).$$

Proof. The corollary follows immediately, as the only difference in the case of good embeddings is that the root of S is always embedded in the root of the underlying tree. Thus, we have to omit the path of left-or-right trees in the beginning. This corresponds to a multiplication by the factor $(1 - 2zB(z))$. □

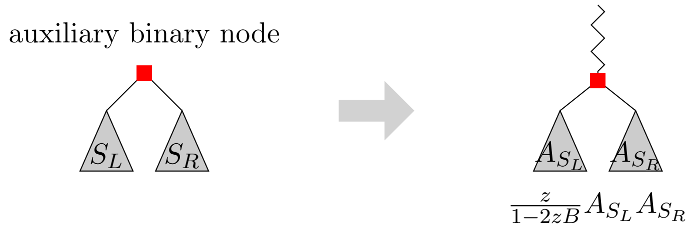


Figure 5: Extension of the recursive construction of the generating function $A_S(z)$ shown in Figure 3 if the auxiliary binary vertex occur. It is embedded according to the picture above and depicted as a red square.

5 Asymptotics of the number of embeddings in $\mathcal{B}_n$

In this section we will do a singularity analysis of the generating functions obtained in the previous section in order to crank out the coefficient asymptotics. Then we will generalize the result to embeddings of disconnected structures.

5.1 Asymptotics of $a_{\mathcal{B}_n}(S)$ and $g_{\mathcal{B}_n}(S)$

Theorem 12. *Consider a rooted tree S with $d_S = (l, u, d_2, \ldots, d_{m-1})$ being its degree distribution sequence. Let $C = \prod_{i=3}^{m-1} (\boldsymbol{C}_{i-1})^{d_i}$. The asymptotics of the number of all embeddings of S into $\mathcal{B}_n$ is given by*

$$a_{\mathcal{B}_n}(S) \sim \frac{C \cdot 2^{\frac{5-m-3l}{2}}}{\Gamma(\frac{m+l-1}{2})} \cdot 2^n \cdot n^{\frac{m+l-3}{2}}$$

for n being odd and $a_{\mathcal{B}_n}(S) = 0$ for n being even. The asymptotics of the number of good embeddings of S into $\mathcal{B}_n$ is given by

$$g_{\mathcal{B}_n}(S) \sim \begin{cases} \dfrac{C \cdot 2^{\frac{6-m-3l}{2}}}{\Gamma(\frac{m+l-2}{2})} \cdot 2^n \cdot n^{\frac{m+l-4}{2}} & \text{if} \quad m+l-2 > 0; \\ \dfrac{\sqrt{2} \cdot 2^n}{\sqrt{\pi n^3}} & \text{if} \quad m+l-2 = 0 \end{cases}$$

for n being odd and $g_{\mathcal{B}_n}(S) = 0$ for n being even.

Proof. Recall that $a_{\mathcal{B}_n}(S) = [z^n]A_S(z)$. The function $A_S(z)$ has two dominant singularities at $\rho_0 = 1/2$ and $\rho_1 = -1/2$, coming from the function $B(z)$ and from the zero in the denominator of the expression for $A_S(z)$ given in (4). Expanding $A_S(z)$ into its Puiseux series for $z \to \rho_0 = 1/2$ gives

$$A_S(z) = C \cdot 2^{\frac{3-m-3l}{2}} \cdot \left(1 - \frac{z}{\rho_0}\right)^{-\frac{m+l-1}{2}} \left(1 + O\left(\left(1 - \frac{z}{\rho_0}\right)^{\frac{1}{2}}\right)\right).$$

Note that $m+l-1 \geqslant 1$, since always $l \geqslant 1$ and $m \geqslant 1$. Expanding $A_S(z)$ into a Puiseux series for $z \to \rho_1 = -1/2$ gives

$$A_S(z) = -C \cdot 2^{\frac{3-m-3l}{2}} \cdot \left(1 - \frac{z}{\rho_1}\right)^{-\frac{m+l-1}{2}} \left(1 + O\left(\left(1 - \frac{z}{\rho_1}\right)^{\frac{1}{2}}\right)\right).$$

By Lemma 1 we get

$$\begin{aligned}[z^n]A_S(z) &\sim \frac{C \cdot 2^{\frac{3-m-3l}{2}}}{\Gamma(\frac{m+l-1}{2})} \cdot (\rho_0)^{-n} \cdot n^{\frac{m+l-3}{2}} - \frac{C \cdot 2^{\frac{3-m-3l}{2}}}{\Gamma(\frac{m+l-1}{2})} \cdot (\rho_1)^{-n} \cdot n^{\frac{m+l-3}{2}} \\ &= \begin{cases} \dfrac{C \cdot 2^{\frac{5-m-3l}{2}}}{\Gamma(\frac{m+l-1}{2})} \cdot 2^n \cdot n^{\frac{m+l-3}{2}} & \text{if } \ n \text{ is odd,} \\ 0 & \text{if } \ n \text{ is even.} \end{cases}\end{aligned}$$

The asymptotic analysis for the number of good embeddings is analogous. Again, $g_{\mathcal{B}_n}(S) = [z^n]G_S(z)$ and $G_S(z)$ has two dominant singularities at $1/2$ and $-1/2$. For $m+l-2 > 0$ we obtain

$$[z^n]G_S(z) \sim \begin{cases} \dfrac{C \cdot 2^{\frac{6-m-3l}{2}}}{\Gamma(\frac{m+l-2}{2})} \cdot 2^n \cdot n^{\frac{m+l-4}{2}} & \text{if } \ n \text{ is odd,} \\ 0 & \text{if } \ n \text{ is even.} \end{cases}$$

The case $m+l-2 = 0$ needs to be treated separately. Note that then $m = 1$ and $l = 1$, thus the structure S that we embed is a single vertex. Therefore the number of good embeddings is just the cardinality of $\mathcal{B}_n$, *i.e.* $g_{\mathcal{B}_n}(S) = \boldsymbol{C}_{\frac{n-1}{2}} \sim \frac{\sqrt{2}\cdot 2^n}{\sqrt{\pi n^3}}$. (Note also that for S being a single vertex we have $a_{\mathcal{B}_n}(S) = n\boldsymbol{C}_{\frac{n-1}{2}} \sim \frac{\sqrt{2}\cdot 2^n}{\sqrt{\pi n}}$.) □

5.2 Embedding disconnected structures in $\mathcal{B}_n$

Now, let us briefly discuss the case of embedding disconnected structures in $\mathcal{B}_n$. Note that in this case all the embeddings must be bad (the underlying structure T has only one maximal element $\mathbb{1}_T$; as long as the induced structure is disconnected, we can be sure that it does not contain the root $\mathbb{1}_T$).

Assume that S is a forest, *i.e.* a set of rooted trees $S_1, S_2, \ldots, S_r$ ($r \geqslant 2$) with the degree distribution sequence $d_S = (l, u, d_2, \ldots, d_{m-1})$. The underlying structure T is connected, thus $S_1, S_2, \ldots, S_r$ always have a common parent in T. Let $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_r)$ be a permutation of the set $\{1, 2, \ldots, r\}$. Define $S^{(\sigma)}$ to be a structure constructed as shown in Figure 6 - we add an additional vertex $\mathbb{1}_{S^{(\sigma)}}$ to S, which is a common parent of $S_1, S_2, \ldots, S_r$ appearing in the order given by σ. Now, instead of counting the number of embeddings of S into T we can simply count the numbers of good embeddings of $S^{(\sigma)}$ in T for all permutations σ generating non-isomorphic structures $S^{(\sigma)}$ and sum them up. Thus,

$$a_{\mathcal{B}_n}(S) = \sum_{\sigma \in \Sigma} g_{\mathcal{B}_n}(S^{(\sigma)}),$$

where Σ is a set of permutations of $\{1, 2, \ldots, r\}$ such that whenever $\sigma, \tau \in \Sigma$ and $\sigma \neq \tau$ then $S^{(\sigma)}$ and $S^{(\tau)}$ are not isomorphic. Moreover, whenever τ is a permutation of $\{1, 2, \ldots, r\}$ and $\tau \notin \Sigma$ then there exists $\sigma \in \Sigma$ such that $S^{(\sigma)}$ and $S^{(\tau)}$ are isomorphic.

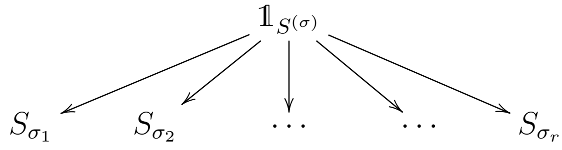


Figure 6: The structure of $S^{(\sigma)}$, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_r)$.

Note that the asymptotics of $g_{\mathcal{B}_n}(S^{(\sigma)})$ is the same for all $\sigma \in \Sigma$ since the degree distribution sequence of $S^{(\sigma)}$ is the same for all $\sigma \in \Sigma$. It is given by $d_{S^{(\sigma)}} = (\tilde{d}_0, \tilde{d}_1, \ldots, \tilde{d}_{m-1}) = (l, u, \ldots, d_{r-1}, d_r + 1, d_{r+1}, \ldots, d_{m-1})$. Therefore, by Theorem 12

$$a_{\mathcal{B}_n}(S) \sim \begin{cases} \dfrac{m!}{k_1! k_2! \cdots k_\ell!} \dfrac{\tilde{C} \cdot 2^{\frac{6-m-3l}{2}}}{\Gamma(\frac{m+l-2}{2})} \cdot 2^n \cdot n^{\frac{m+l-4}{2}} & \text{if} \quad n \text{ is odd,} \\ 0 & \text{if} \quad n \text{ is even} \end{cases}$$

where ℓ is the number of equivalence classes of the set $\{S_1, S_2, \ldots, S_r\}$ with respect to the equivalence relation of being isomorphic and $k_1, k_2, \ldots, k_\ell$ are the cardinalities of those classes. Here $\tilde{C} = \prod_{i=3}^{m-1} (\boldsymbol{C}_{i-1})^{\tilde{d}_i}$. (Note that here we do not consider the case $m+l-2 = 0$ from Theorem 12, because by $r \geqslant 2$ we always have $m + l - 2 > 0$.)

6 Non-plane case - embeddings in $\mathcal{V}_n$

6.1 Embedding rooted trees in $\mathcal{V}_n$

In this subsection we explain how to take advantage of the results obtained for the plane case in order to infer about the asymptotics of good and all embeddings of a rooted tree S in the family of non-plane binary trees $\mathcal{V}_n$.

Theorem 13. *Consider a rooted tree S with $d_S = (l, u, d_2, \ldots, d_{m-1})$ being its degree distribution sequence. The generating function $A_S(z)$ of the sequence $a_{\mathcal{V}_n}(S)$, counting the number of all embeddings of S into the family $\mathcal{V}_n$, is given by*

$$A_S(z) = \left(\frac{1}{1 - zV(z)} \right)^{m+l-1} z^{l+u-1} \; V(z)^{l+u} \; C_S \; (1 + o(1)), \qquad \text{as } z \to \pm\rho, \qquad (7)$$

where C_S is a constant dependent on the structure of S, $V(z)$ is the generating function of the family of non-plane binary trees, cf. (3), *and $\rho \approx 0.6346$ is the radius of convergence of $V(z)$.*

Remark 14. It is well known that $V(z)$ has two dominant singularities, $\pm\rho$, a result going back to Otter [39]. Note that the value for ρ stated above does not coincide with the one reported in [16, Chapter VII], since we use different size functions: There the size of a tree was defined as the number of internal nodes, while we count the total number of nodes. This is the reason why in our model the coefficients $v_n = [z^n]V(z)$ are zero for even n, which yields a periodicity in the generating function that results in the presence of two dominant singularities. However, the generating function $N(z)$ of non-plane binary trees where z solely marks the number of internal vertices can easily be connected with our generating function $V(z)$ via $V(z) = zN(z^2)$. Thus, with the result from [16] that

$$N(z) \sim \frac{1}{\sigma} - a\sqrt{1 - \frac{z}{\sigma}}, \qquad \text{as } z \to \sigma$$

with $\sigma \approx 0.4027$ and $a \approx 2.8062$, we immediately know that there are two dominant singularities of $V(z) = zN(z^2)$ at $z = \pm\sqrt{\sigma}$ and we get

$$V(z) = zN(z^2) \sim \pm\sqrt{\sigma}\left(\frac{1}{\sigma} - a\sqrt{2}\sqrt{1 \mp \frac{z}{\sqrt{\sigma}}}\right), \qquad \text{as } z \to \pm\sqrt{\sigma}.$$

Finally, by setting $\rho = \sqrt{\sigma} \approx 0.6346$ and $b = a\sqrt{2\sigma} \approx 2.5184$ we obtain

$$V(z) \sim \pm\left(\frac{1}{\rho} - b\sqrt{1 \mp \frac{z}{\rho}}\right), \qquad \text{as } z \to \pm\rho.$$

Proof. This time we introduce a bivariate generating function, where z still marks the total number of vertices of a tree, while u is associated with *classes of vertices*. Two vertices v, w are meant to belong to the same class whenever there exists an isomorphism $f : T \to T$ such that $f(v) = w$. From [35] we have

$$V(z,u) = zu + \frac{zu}{2}(V(z,u)^2 - V(z^2,u^2) + 2V(z^2,u)). \tag{8}$$

This can be seen from specifying $\mathcal{V}$ in a different way:

$$\mathcal{V} = \{\bullet\} + \{\bullet\} \times \mathcal{S}et_2(\mathcal{V}) + \{\bullet\} \times \Delta(\mathcal{V} \times \mathcal{V}).$$

Indeed, if the subtrees of the root are non-isomorphic, then we simply inherit the vertex classes from the subtrees. Translating into generating function (see the set construction presented in Section 2 and Equation (1)) gives the first two summands of (8). Otherwise, the subtrees of the root are identical. This doubles the number of vertices, but each vertex is in the same vertex class as its duplicate. This yields the term $V(z^2, u)$.

By $V_u(z,u)$ we denote the derivative of $V(z,u)$ with respect to u, *i.e.* $V_u(z,u) = \frac{\partial V(z,u)}{\partial u}$. What we will need is $V_u(z,1)$. Differentiating Equation (8) with respect to u and plugging $u = 1$ yields

$$V_u(z,1) = \frac{V(z)}{1 - zV(z)}. \tag{9}$$

Case 1: S is a Motzkin tree.

We proceed as in the plane case and start with recursively defining the generating function $A_S(z)$ for the number of embeddings of S into the family $\mathcal{V}_n$, when S is a Motzkin tree. As in the plane case, we denote the single subtree attached to the root by $\tilde{S}$ in case of a unary root and the two subtrees by S_L and S_R (arbitrarily chosen order) in case of a binary root. We obtain

$$A_S(z) = \begin{cases} V_u(z,1) & \text{if } S = \bullet, \\ \dfrac{zV(z)}{1-zV(z)} A_{\tilde{S}}(z) & \text{if } S = (\bullet, \tilde{S}), \\ \dfrac{z}{(1-zV(z))^2} A_{S_L}(z) A_{S_R}(z) & \text{if } S = (\bullet, S_L, S_R) \quad \text{and} \quad S_L \ncong S_R, \\ \dfrac{z}{(1-zV(z))^2} \dfrac{1}{2}(A_{S_L}(z)^2 + A_{S_L}(z^2)) & \text{if } \mathcal{S} = (\bullet, S_L, S_R) \quad \text{and} \quad S_L \cong S_R. \end{cases}$$

The idea of setting up this recursive definition for $A_S(z)$ is similar to the plane case with the following differences. In the first case, corresponding to embedding a single node, we can mark an arbitrary vertex class, instead of an arbitrary vertex, since there might be some non-trivial isomorphisms that would lead to multiple countings of the same embedding. Furthermore, the paths of left-or-right trees from the previous section, yielding a factor $1/(1-2zB(z))$, are now replaced by paths of trees where we do not distinguish between the left-or-right order, since we are in the non-plane setting. Thus, these paths give a factor $1/(1-zV(z))$. Finally, in the case when the Motzkin tree starts with a binary root, we have to distinguish between the cases whether the two attached trees are isomorphic or not. The non-isomorphic case works analogously to its plane version, while in the isomorphic case we have to eliminate potential double-countings by using the same idea as for Equation (3) (compare also with the multiset construction presented in Section 2). We do not have to solve the recursion for $A_S(z)$ explicitly, since we are solely interested in the asymptotic behaviour of its coefficients and it is easy to see that asymptotically the contribution of the term $A_{S_L}(z^2)$ is negligible. Indeed, since $\rho < 1$ the function $A_S(z^2)$ is analytic at $z = \rho$. Thus, $[z^n]A_S(z^2) < (\rho + \varepsilon)^{-n}$, which is exponentially smaller than $C\rho^{-n}n^{\beta} = [z^n]A_S(z)$.

Thus, by iterating we obtain

$$A_S(z) \sim \left(\frac{z}{(1-zV(z))^2}\right)^{l-1} V_u(z,1)^l \left(\frac{zV(z)}{1-zV(z)}\right)^u \left(\frac{1}{2}\right)^s, \qquad \text{as } z \to \rho,$$

where l denotes the number of leaves, u the number of unary nodes and s the number of symmetry nodes in S (a symmetry node is a parent of two isomorphic subtrees). The same expansion holds for $z \to -\rho$.

Case 2: S contains vertices with more than two children.

In the general case where S is an arbitrary non-plane tree, *i.e.* a Pólya tree, we proceed as in the previous section and consider the embeddings of all non-plane unary-binary trees obtained by replacing d-ary nodes with $d \geqslant 3$ together with their children by binary trees with d leaves. Thus, again taking into account that there are $m-l-u$ binary nodes that were already there before the replacement (as binary or d-ary nodes with $d \geqslant 3$) and $2l+u-m-1$ auxiliary binary nodes that were introduced by the replacement, we get

$$A_S(z) \sim \left(\frac{z}{1-zV(z)}\right)^{2l+u-m-1} \left(\frac{z}{(1-zV(z))^2}\right)^{m-l-u} V_u(z,1)^l \left(\frac{zV(z)}{1-zV(z)}\right)^u C_S, \quad \text{as } z \to \rho, \tag{10}$$

and the analogous expansion for $z \to -\rho$. The constant C_S arises from the isomorphisms and reads as

$$C_S = \sum_{\substack{t \in \mathcal{M}_S \\ s \text{ symmetry node of } t}} \left(\frac{1}{2}\right)^s, \tag{11}$$

where $\mathcal{M}_S$ denotes the set of all non-plane unary-binary trees obtained from S by replacing the d-ary nodes with non-plane binary trees with d leaves for $d \geqslant 3$.

Finally, substituting the expression given in (9) for $V_u(z,1)$ in Equation (10) yields the desired result. Note that the asymptotic equivalence (10), or (7) respectively, is also true for the case when S is a single node, *i.e.* $l=1$ and $u=s=0$. □

Theorem 15. *Consider a rooted tree S with $d_S = (l, u, d_2, \ldots, d_{m-1})$ being its degree distribution sequence. The asymptotics of the number of all embeddings of S into $\mathcal{V}_n$ is given by*

$$a_{\mathcal{V}_n}(S) \sim \frac{2C_S b^{-m-l+1}\rho^{-m-l}}{\Gamma(\frac{m+l-1}{2})} \cdot \rho^{-n} \cdot n^{\frac{m+l-3}{2}}$$

for n being odd and $a_{\mathcal{V}_n}(S) = 0$ for n being even. The asymptotics of the number of good embeddings of S into $\mathcal{V}_n$ is given by

$$g_{\mathcal{V}_n}(S) \sim \begin{cases} \dfrac{2C_S b^{-m-l+2}\rho^{-m-l+1}}{\Gamma(\frac{m+l-2}{2})} \cdot \rho^{-n} \cdot n^{\frac{m+l-4}{2}} & \text{if } \ m+l-2 > 0 \\ \dfrac{b}{\sqrt{\pi}} \cdot \rho^{-n} \cdot n^{-3/2} & \text{if } \ m+l-2 = 0, \end{cases}$$

for n being odd and $g_{\mathcal{V}_n}(S) = 0$ for n being even. Here $b \approx 2.5184$, $\rho \approx 0.6346$ and the constant C_S, given in (11), depends on the structure of S.

Proof. First, note that $V(\rho) \sim \frac{1}{\rho}$, which was already outlined in Remark 14. Therefore, the dominant part of the asymptotics of the coefficients of $A_S(z)$ comes from the factors $1/(1-zV(z))$, which give

$$\frac{1}{1-zV(z)} \sim \frac{1}{\rho b\sqrt{1-\frac{z}{\rho}}} \qquad \text{for } z \to \rho.$$

The result for $a_{\mathcal{V}_n}(S)$ follows immediately by use of Lemma 1. As in the plane case, the generating function $G_S(z)$ for the good embeddings just differs from $A_S(z)$ by a factor $(1-zV(z))$ and thus, the asymptotic behaviour of its coefficients can be determined analogously. Recall that $m+l-2=0$ represents the case where S is a single vertex. The number of good embeddings is therefore just the cardinality of $\mathcal{V}_n$ (see Remark 14). □

6.2 Embedding disconnected structures in $\mathcal{V}_n$

Now, let us comment on embedding disconnected structures in a non-plane case. Let S be a forest, *i.e.* a set of rooted trees $S_1, S_2, \ldots, S_r$, $r \geqslant 2$. Again, instead of counting all embeddings of S into $\mathcal{V}_n$, we can count the good embeddings of $\tilde{S}$ in $\mathcal{V}_n$, where $\tilde{S}$ is a forest S with an additional common parent that clips together all S_i's. Note that in the non-plane case the order of S_i's does not matter, thus we simply have

$$a_{\mathcal{V}_n}(S) = g_{\mathcal{V}_n}(\tilde{S}).$$

7 Planted plane case - embeddings in $\mathcal{T}_n$

In this section we extend the results from plane binary trees to planted plane trees, *i.e.* to rooted trees where each internal node can have arbitrarily many child-nodes and the order of the subtrees is important. The structures that we embed are as well planted plane trees, and therefore every such a tree S is of the form $S = (\bullet, S_1, \ldots, S_k)$, where the S_i's denote the subtrees that are attached to the root. We will present a lemma containing the construction of the generating function $A_S(z)$ of all embeddings of the tree S in the family $\mathcal{T}_n$ of planted plane trees of size n. But before doing so, we state another lemma that lists all the building blocks of that construction together with their generating functions.

Lemma 16. *The following constructions made from the class $\mathcal{T}$ are needed in the sequel:*

The generating function associated with sequences of planted plane trees (plane forests), denoted by $\mathcal{S}eq(\mathcal{T})$, is $z/(1 - T(z))$.

Let $\mathcal{P}$ denote the class of all paths of the following form: Take a path in the graph theoretical sense (chain of vertices) as a spine and attach to each of its vertices a sequence of planted plane trees left of the spine (left forest) and a sequence of planted plane trees right of the spine (right forest). Then remove the last vertex, say v, of the spine, but keep its edge as a separator between the left and the right forest attached to the parent node of x. Then the generating function of $\mathcal{P}$ is

$$P(z) = \frac{1 - T(z)}{1 - 2T(z)}. \tag{12}$$

The generating functions $G_S(z)$ and $B_S(z)$ associated with good and bad embeddings, respectively, of S into T satisfy

$$G_S(z) = \frac{1 - 2T(z)}{1 - T(z)} A_S(z) \quad \text{and} \quad B_S(z) = \frac{T(z)}{1 - T(z)} A_S(z). \tag{13}$$

Proof. The first assertion was already presented in Section 2 and only listed for the sake of completeness.

To prove the assertion on $\mathcal{P}$, observe that $\mathcal{P}$ is built of segments made of the left forest, the spine node and the right forest. So, we have actually a sequence of such segments. Thus

$$\mathcal{P} = \mathcal{S}eq(\mathcal{S}eq(\mathcal{T}) \times \{\bullet\} \times \mathcal{S}eq(\mathcal{T})).$$

Translating into generating functions gives

$$P(z) = \frac{1}{1 - \frac{z}{(1-T(z))^2}} = \frac{1-T(z)}{1-T(z)-\frac{z}{1-T(z)}} = \frac{1-T(z)}{1-2T(z)},$$

where we used the functional equation $T(z) = z/(1-T(z))$ in the last step.

A general embedding can be seen as a path of type $\mathcal{P}$ from the root of T to the embedded root of S, call it v, and the subtree rooted at v. This subtree can in turn be interpreted as a new tree with a good embedding of S. Thus $A_S(z) = P(z)G_S(z)$, as desired. Finally, $B_S(z) = A_S(z) - G_S(z)$. □

Lemma 17. *The generating function $A_S(z)$ of all embeddings of $S = (\bullet, S_1, \ldots, S_k)$ into the family $\mathcal{T}_n$ of planted plane trees of size n can be recursively specified as*

$$A_S(z) = \begin{cases} zT'(z) = \dfrac{T(z)(1-T(z))}{1-2T(z)} & \text{if } k=0, \\[2ex] \dfrac{T(z)}{1-2T(z)} A_{S_1}(z) & \text{if } k=1, \\[2ex] \dfrac{T(z)^2}{(1-2T(z))^2(1-T(z))} A_{S_1}(z)A_{S_2}(z) & \text{if } k=2, \\[2ex] \dfrac{T(z)}{(1-2T(z))^2}\Bigg(\dfrac{1-2T(z)}{1-T(z)} A_{S_1}(z)A_{S_{2,k}}(z) & \\ \quad +\dfrac{T(z)(1-2T(z))}{(1-T(z))^2} A_{S_{1,k-1}}(z)A_{S_k}(z) & \\ \quad + \left(\dfrac{1-2T(z)}{1-T(z)}\right)^2 \left(A_{S_{1,2}}(z)A_{S_{3,k}}(z) + \cdots + A_{S_{1,k-2}}(z)A_{S_{k-1,k}}(z)\right)\Bigg) & \text{if } k>2, \end{cases} \tag{14}$$

where $T(z)$ denotes the generating function of the family of planted plane trees and $S_{i,j}$ denotes the tree $S_{i,j} = (\bullet, S_i, \ldots, S_j)$ that consists of a root to which the $j-i+1$ subtrees $S_i, \ldots, S_j$ are attached (in that order).

Proof. The case $k=0$ is equivalent to the binary cases, and corresponds to marking an arbitrary node in the tree T. Differentiating both sides of the specification $T(z) = z/(1-T(z))$ of planted plane trees with respect to z and solving for $T'(z)$ yields the equality

$$zT'(z) = \frac{z}{1-2T(z)} = \frac{T(z)(1-T(z))}{1-2T(z)}.$$

In the case $k=1$ we have to embed the root of S at some node v and then its single subtree S_1 into one of the subtrees attached to v in T. Figure 7 visualizes such an embedding. The generating function of this particular subtree is $A_{S_1}(z)$. To the left and to the right of this subtree we have a sequence of planted plane trees dangling from v.

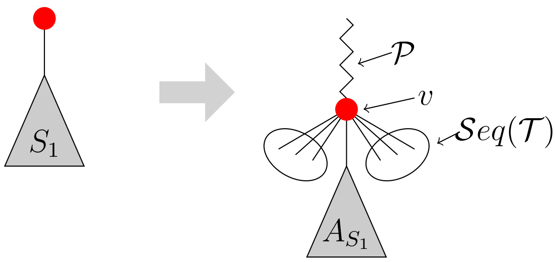


Figure 7: Sketch of the principle of embedding a plane tree $S = (\bullet, S_1)$ into the family of planted plane trees (case $k = 1$).

From the root to v there is a path of type $\mathcal{P}$, the class presented in Lemma 16. Altogether, this gives

$$A_S(z) = P(z)\frac{z}{(1-T(z))^2}A_{S_1}(z).$$

From $z = T(z)(1 - T(z))$ and (12) we get the expression listed in (14) after all.

The obvious generalization to the case $k = 2$ is as follows: Embed the root of S into a vertex v of T. When embedding S_1 and S_2, there will be a last common ancestor of the roots of S_1 and S_2 until which the paths to the two subpatterns coincide and split into two paths afterwards. Call this "splitting node" w. Then we have paths of type $\mathcal{P}$ from the root (of T) to v and from v to w, together contributing a factor $P(z)^2$. Moreover, v and w themselves contribute each a factor z and there is again a left forest and a right forest dangling from v and likewise from w. At the splitting node, we have a left forest, a middle forest and a right forest, separated by the two subtrees carrying the embeddings of S_1 and S_2. Altogether, we obtain

$$A_S(z) = P(z)^2\frac{z^2}{(1-T)^5}A_{S_1}A_{S_2} \tag{15}$$

which simplifies to the expression listed in (14). Figure 8 visualizes the described approach.

The disadvantage of this approach is that in the general case the subtrees of the splitting node may contain embeddings of more than one subtree of S, and we would have to distinguish all cases induced by the possible partitions of the subtrees of S into the subtrees of the splitting node. Hence, let us offer a second approach to the case $k = 2$ by decomposing the structure in another way, which allows for easy generalization: We leave the two paths of type $\mathcal{P}$, the two forests dangling from v and the left forest dangling from w as in the above construction. But then we decompose what remains into the subtree containing the embedding of S_1 (blue tree in Figure 9) and the remaining structure (green tree in Figure 9). This remaining structure, which consists actually of the splitting node, the middle and the right forest and the subtree with the embedding of S_2 in the former decomposition, is a tree rooted at w and containing the embedding of S_2. Therefore it is in fact a tree with a bad embedding of S_2. Indeed, the root of S_2 cannot be embedded at the splitting node, as then we would have chosen the parent of w as splitting node. With

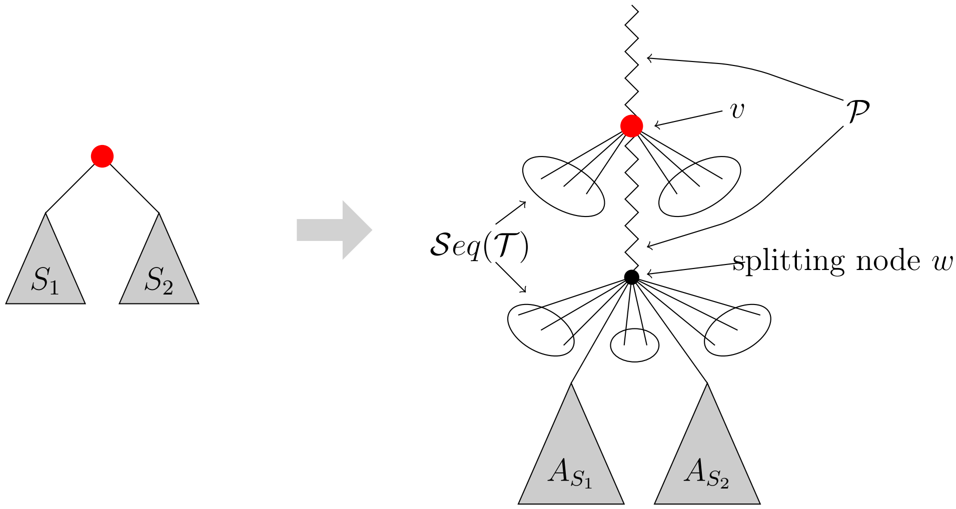


Figure 8: Sketch of the principle of embedding a plane tree $S = (\bullet, S_1, S_2)$ into the family of planted plane trees (case $k = 2$, first approach).

this we get

$$A_S(z) = P(z)^2 \frac{z}{(1-T)^3} A_{S_1} B_{S_2} = P(z)^2 \frac{z}{(1-T)^3} \frac{T(z)}{(1-T(z))} A_{S_1} A_{S_2},$$

which is of course the same as (15).

Now, let us continue with the proof of the recurrence for the case $k > 2$. In order to do so let us look at Figure 10 that visualizes how an embedding of a tree S in a tree T can be constructed. Again we keep everything up to the splitting node, namely the two paths of type $\mathcal{P}$, the node v where the root of S is embedded, the left and right forests attached to v and the left forest attached to the splitting node, altogether yielding $P(z)^2 z/(1-T(z))^3 = T(z)/(1-2T(z))^2$ as a universal factor for all the subcases discussed below.

As before, we call the subtree that is attached to the splitting and contains the embedding of S_1 the "blue subtree" and the remaining structure the "green subtree", which is a tree rooted at w. Note that the blue subtree may also contain embeddings of further subtrees of S. If it also contains an embedding of S_i, then planarity implies that $S_2, \ldots S_{i-1}$ are embedded into the blue subtree as well. Therefore we have the following cases:

- Solely S_1 is embedded in the blue subtree. Then the green subtree contains the embeddings of $S_2, \ldots, S_k$, which is equivalent to embed $S_{2,k} = (\bullet, S_2, \ldots, S_k)$, as long as the root of $S_{2,k}$ is embedded into the splitting node in order to prevent multiple counting. So, in this case the blue subtree is a tree that contains any (good or bad) embedding of S_1, giving a factor $A_{S_1}(z)$, while the green subtree can only contain a good embedding of $S_{2,k}$, contributing a factor $\frac{1-2T(z)}{1-T(z)} A_{S_{2,k}}(z)$ by Lemma 16.

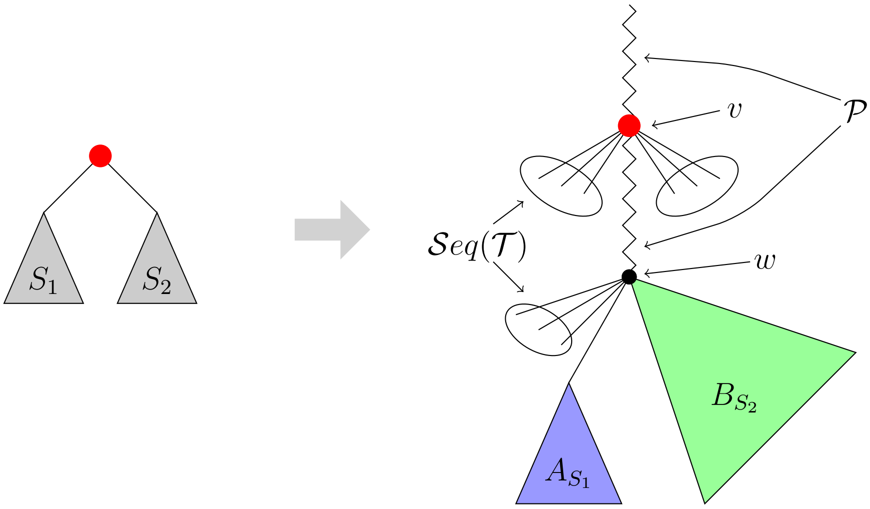


Figure 9: Sketch of the principle of embedding a plane tree $S = (\bullet, S_1, S_2)$ into the family of planted plane trees (case $k = 2$, second approach).

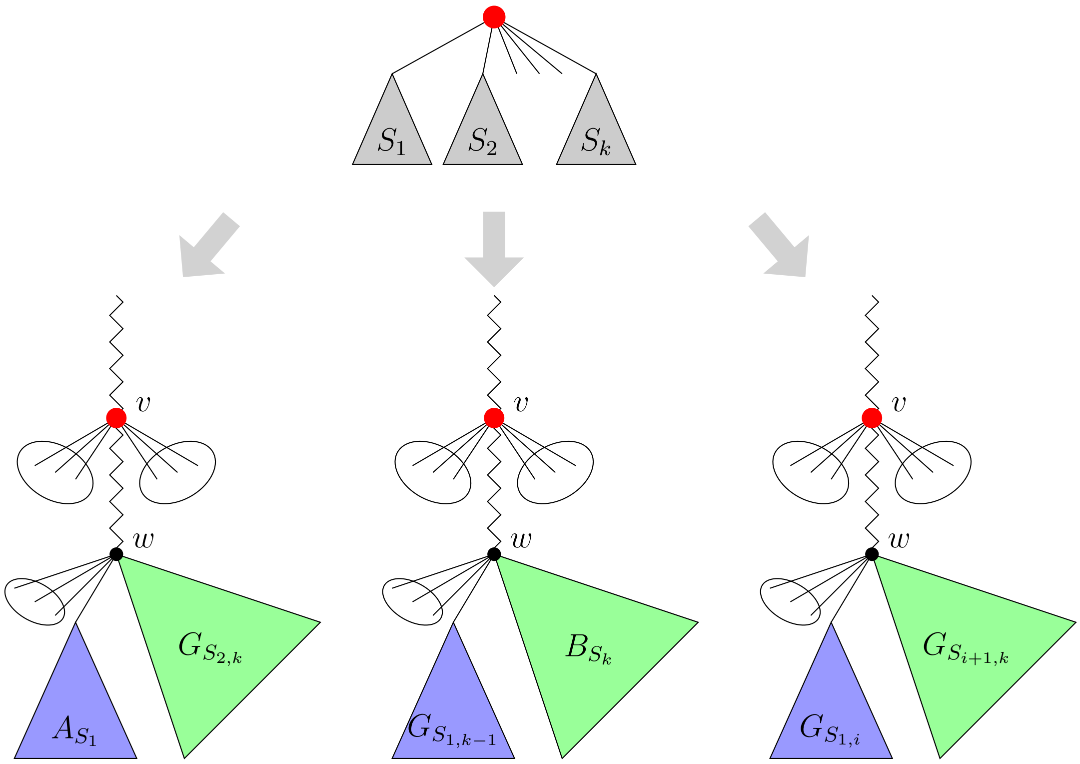


Figure 10: Sketch of the principle of embedding a plane tree $S = (\bullet, S_1, \ldots, S_k)$ (for $k > 2$) into the family of planted plane trees. We distinguish three sub-cases: (*left*) solely S_1 is embedded in the blue subtree, (*middle*) solely S_k is embedded in the green subtree, (*right*) neither the blue nor the green subtree contain embeddings of only one of the subtrees $S_1, \ldots, S_k$.

- Solely S_k is embedded in the green subtree. Here we count the good embeddings of $S_{1,k-1}$ in the blue subtree, as this is necessary for all cases where we consider more than just one of the S_i's to be embedded in the same subtree. However, in this case we have to count only the bad embeddings of S_k in the green subtree, since no node of S (except the root of S) can be embedded into the splitting node (compare with the case $k = 2$ above). Altogether this yields the factor $\frac{1-2T(z)}{1-T(z)} \frac{T(z)}{1-T(z)} A_{S_{1,k-1}}(z) A_{S_k}(z)$.

- Neither the blue nor the green subtree contain embeddings of only one of the subtrees $S_1, \ldots, S_k$. Then we face a good embedding of $S_{1,i}$ into the blue subtree and a good embedding of $S_{i+1,k}$ into the green subtree, yielding together a factor $\left(\frac{1-2T(z)}{1-T(z)}\right)^2 A_{S_{1,i}}(z) A_{S_{i+1,k}}(z)$.

Together with the universal factor from above we get the desired coefficients. □

Remark 18. Note that for the cases $k = 0, 1, 2$ the generating function $A_S(z)$ of all embeddings of $S = (\bullet, S_1, \ldots, S_k)$ into the family $\mathcal{T}_n$ of planted planes trees of size n given in (14) is of the form $f_k(T) \cdot A_{S_1}(z) \cdots A_{S_k}(z)$, where $f_k(T)$ is a function that depends only on $T(z)$ and on the size k of S, but not on the specific shape of S. We want to emphasize that, by digging into the structure of S and by recursive application of the formulas given in (14), it follows that $A_S(z)$ is in fact of the form

$$A_S(z) = f(T) \cdot A_{S_1}(z) \cdots A_{S_k}(z),$$

for arbitrary $S = (\bullet, S_1, \ldots, S_k)$.

Now, we are in the position to obtain the asymptotic number of all and good embeddings of a given plane tree S in the family of planted plane trees.

Theorem 19. *Consider a rooted tree S of size m with degree distribution sequence $d_S = (l, d_1, d_2, \ldots, d_{m-1})$. Let $C = \prod_{i=1}^{m-1} (\boldsymbol{C}_{i-1})^{d_i}$. The asymptotics of the number of all embeddings of S into $\mathcal{T}_n$ is given by*

$$a_{\mathcal{T}_n}(S) \sim \frac{C \cdot (\frac{1}{2})^{m+l}}{\Gamma(\frac{m+l-1}{2})} \cdot 4^n \cdot n^{\frac{m+l-3}{2}}.$$

The asymptotics of the number of good embeddings of S into $\mathcal{T}_n$ is given by

$$g_{\mathcal{T}_n}(S) \sim \frac{2C \cdot (\frac{1}{2})^{m+l}}{\Gamma(\frac{m+l-2}{2})} \cdot 4^n \cdot n^{\frac{m+l-4}{2}}.$$

Proof. Triggered by the observation in Remark 18, let us set

$$f_1(z) = \frac{1}{2(1-2T(z))}, \quad \text{and} \quad f_k(z) = \frac{A_S(z)}{\prod_{i=1}^{k} A_{S_i}(z)} \quad \text{for } k > 1. \tag{16}$$

Then (14) immediately gives $f_2(z) = T(z)^2/((1-2T(z))^2(1-T(z)))$.

Next, consider the last equation of (14) (the case $k \geqslant 3$) and observe that all generating functions on the right-hand side which are associated with a composite structure are of the form $A_{S_{i,j}}(z)$, where the root of $S_{i,j}$ has degree at least two. Thus, dividing the equation by $\prod_{i=1}^{k} A_{S_i}(z)$ (and cancelling out all single $A_{S_j}(z)$) yields only quotients which can be readily turned into $f_\ell(z)$ with suitable choices of ℓ, because the case $\ell = 1$ does not appear here. A straight-forward simplification then gives

$$f_k(z) = \frac{T(z)}{(1-T(z))^2} \sum_{j=1}^{k-1} f_j(z) f_{k-j}(z) \quad \text{for } k \geqslant 3. \tag{17}$$

Both sides of this equation tend to infinity, as $z \to 1/4$, and we need their singular behaviour for our analysis of $A_S(z)$. Hence, we set $g_k(z) = (1-2T(z))^k f_k(z)$ for $k \geqslant 1$. Plugging this into (17) we observe that $g_k(z)$ satisfies the same recurrence as $f_k(z)$, but with the initial values $g_1(z) = 1/2$ and $g_2(z) = T(z)^2/(1-T(z))$. As $T(1/4) = 1/2$, the functions $g_k(z)$ are finite at $z = 1/4$. By evaluating the recurrence at $z = 1/4$ and setting $h_k := 2g_k(1/4)$, we get a recurrence for h_k, which is in fact already valid for $k \geqslant 2$:

$$h_1 = 1 \quad \text{and} \quad h_k = \sum_{j=1}^{k-1} h_j h_{k-j} \quad \text{for } k \geqslant 2.$$

This is exactly the recurrence for the Catalan numbers, and thus, $h_k = \boldsymbol{C}_{k-1}$.

Hence, for $z \to 1/4$ and $k \geqslant 2$ we have

$$f_k(z) \sim \frac{1}{2} \boldsymbol{C}_{k-1} (1-4z)^{-k/2},$$

which implies that as $z \to 1/4$ we have

$$A_S(z) \sim \frac{\boldsymbol{C}_{k-1}}{2} (1-4z)^{-k/2} A_{S_1}(z) \cdots A_{S_k}(z) = \left(\prod_{i=1}^{m-1} \left(\frac{\boldsymbol{C}_{i-1}}{2} (1-4z)^{-i/2} \right)^{d_i} \right) (A_\bullet(z))^l,$$

where $S = (\bullet, S_1, \ldots, S_k)$, d_i denotes the number of nodes with out-degree i, l denotes the number of leaves, *i.e.* $l = d_0$, and the equation follows from recursively going into the subtrees $S_1, \ldots, S_k$ and using (16) until one encounters a leaf of S. Then each leaf yields a factor $A_\bullet(z)$. Using the equality $A_\bullet(z) = zT'(z) \sim \frac{1}{2}(1-4z)^{-1/2}$, which follows from (14) and the singular expansion of $T(z)$, we get for $z \to \frac{1}{4}$

$$A_S(z) \sim \left(\prod_{i=1}^{m-1} \left(\frac{\boldsymbol{C}_{i-1}}{2} \right)^{d_i} \right) (1-4z)^{-\left(l + \sum_{i=1}^{m-1} i d_i\right)/2} \left(\frac{1}{4} \right)^l. \tag{18}$$

Note that $\sum_{i=1}^{m-1} i d_i = m - 1$, since every vertex with out-degree i is counted exactly i times and thus, we simply obtain the total number of nodes with in-degree greater than zero (*i.e.* all nodes except for the root). We also have $\sum_{i=1}^{m-1} d_i = m - l$ thus (recall that $C = \prod_{i=1}^{m-1} (\boldsymbol{C}_{i-1})^{d_i}$)

$$\left(\frac{1}{4}\right)^l \cdot \prod_{i=1}^{m-1} \left(\frac{\boldsymbol{C}_{i-1}}{2}\right)^{d_i} = C \left(\frac{1}{2}\right)^{m+l}.$$

Finally, Lemma 1 gives

$$a_{\mathcal{T}_n}(S) \sim \frac{C \cdot (\frac{1}{2})^{m+l}}{\Gamma(\frac{m+l-1}{2})} \cdot 4^n \cdot n^{\frac{m+l-3}{2}}.$$

The generating function of the number of good embeddings can be derived from the generating function $A_S(z)$ by multiplication by the factor $\frac{1-2T(z)}{1-T(z)}$. This factor is responsible for getting rid of the path of trees which could appear above embedded root of S when we were considering all embeddings. Thus we have $G_S(z) = \frac{1-2T(z)}{1-T(z)} A_S(z)$. Noticing that

$$\frac{1-2T(z)}{1-T(z)} = 2\frac{\sqrt{1-4z}}{1+\sqrt{1-4z}},$$

using (18) and applying Lemma 1 yields the desired result. □

8 Asymptotics and monotonicity of the ratio $g_{\mathcal{F}}(S)/a_{\mathcal{F}}(S)$

Kubicki *et al.* [29] proved that if T is a complete balanced binary tree of arbitrary size and S_1, S_2 are rooted trees in which each node has at most 2 descendants (*i.e.* S_1 and S_2 are Motzkin trees) and $S_1 \subseteq S_2$, then $\frac{g_T(S_1)}{a_T(S_1)} \leqslant \frac{g_T(S_2)}{a_T(S_2)}$. (From now on this property will be interchangeably called "the ratio $\frac{g_T(S)}{a_T(S)}$ being weakly increasing with S".) They also conjectured that the inequality remains true for any rooted trees S_1 and S_2. One year later in [30] they also stated an asymptotic result for the ratio $\frac{g_T(S)}{a_T(S)}$ when S is an arbitrary rooted tree and T a complete binary tree of size n. They showed that $\lim_{n\to\infty} \frac{g_T(S)}{a_T(S)} = 2^{l-1} - 1$ where l is the number of leaves in S. Thereby they proved that for any rooted tree S the asymptotic ratio $\frac{g_T(S)}{a_T(S)}$ is non-decreasing with S (the function $2^{l-1} - 1$ increases with l and if $S_1 \subseteq S_2$ then the number of leaves of S_2 equals at least the number of leaves of S_1).

The conjecture from [29] was disproved by Georgiou [22] who chose specific ternary trees as embedded structures to construct a counterexample. He also generalized the underlying structure to a complete k-ary tree and considered strict-order preserving maps instead of embeddings. In this setting he proved that a correlation inequality (corresponding to $\frac{g_{\mathcal{T}_n}(S_1)}{a_{\mathcal{T}_n}(S_1)} \leqslant \frac{g_{\mathcal{T}_n}(S_2)}{a_{\mathcal{T}_n}(S_2)}$) already holds for S_1, S_2 being arbitrary rooted trees such that $S_1 \subseteq S_2$.

Referring to the asymptotic result from [30], we show that in our case the asymptotic ratios $\frac{\sqrt{n}\, g_{\mathcal{B}_n}(S)}{a_{\mathcal{B}_n}(S)}$, $\frac{\sqrt{n}\, g_{\mathcal{T}_n}(S)}{a_{\mathcal{T}_n}(S)}$ and $\frac{\sqrt{n}\, g_{\mathcal{V}_n}(S)}{a_{\mathcal{V}_n}(S)}$ are all weakly increasing with S for S being an arbitrary rooted tree. Using this asymptotic result we show later that also the ratios

$\frac{g_{\mathcal{B}_n}(S)}{a_{\mathcal{B}_n}(S)}$, $\frac{g_{\mathcal{T}_n}(S)}{a_{\mathcal{T}_n}(S)}$ and $\frac{g_{\mathcal{V}_n}(S)}{a_{\mathcal{V}_n}(S)}$ (unlike in the case from [29]) are eventually weakly increasing with S for sufficiently large n.

We start with the asymptotics of the ratio of the number of good to the number of all embeddings of S into any considered family.

Corollary 20. *Consider a rooted tree S with $d_S = (l, u, d_2, \ldots, d_{m-1})$ being its degree distribution sequence. Let $k = \frac{m+l-2}{2}$ and let $\mathcal{F}_n$ denote any of the families $\mathcal{B}_n$, $\mathcal{V}_n$, $\mathcal{T}_n$. The asymptotic ratio of the number of good embeddings of S into $\mathcal{F}_n$ to the number of all embeddings into $\mathcal{F}_n$ is given by*

$$\frac{g_{\mathcal{F}_n}(S)}{a_{\mathcal{F}_n}(S)} \sim \begin{cases} \dfrac{\Gamma(k+1/2)}{\Gamma(k)} \dfrac{c_{\mathcal{F}_n}}{\sqrt{n}} & \text{if} \quad k > 0, \\ 1/n & \text{if} \quad k = 0, \end{cases}$$

where

$$c_{\mathcal{B}_n} = \sqrt{2}, \qquad c_{\mathcal{V}_n} = b\rho \qquad \text{and} \qquad c_{\mathcal{T}_n} = 2$$

(b and ρ are as in Theorem 15). For the families $\mathcal{B}_n$ and $\mathcal{V}_n$ we consider only n's being odd.

Proof. For $\mathcal{B}_n$, $\mathcal{V}_n$ and $\mathcal{T}_n$ the corollary follows immediately from Theorems 12, 15 and 19, respectively. □

Theorem 21. *Let S_1, S_2 be rooted trees such that $S_1 \subseteq S_2$ and let $\mathcal{F}_n$ denote any of the families $\mathcal{B}_n$, $\mathcal{V}_n$, $\mathcal{T}_n$. Then*

$$\lim_{n\to\infty} \sqrt{n}\, \frac{g_{\mathcal{F}_n}(S_1)}{a_{\mathcal{F}_n}(S_1)} \leqslant \lim_{n\to\infty} \sqrt{n}\, \frac{g_{\mathcal{F}_n}(S_2)}{a_{\mathcal{F}_n}(S_2)}.$$

Proof. Let $d_{S_1} = (l_1, u_1, \ldots)$, $d_{S_2} = (l_2, u_2, \ldots)$, $k_1 = \frac{m_1+l_1-2}{2}$, $k_2 = \frac{m_2+l_2-2}{2}$ (where m_i denotes the size of S_i) and $k_1 > 0$ (the case when $k_1 = 0$ is trivial). By Corollary 20 we have

$$\lim_{n\to\infty} \sqrt{n}\, \frac{g_{\mathcal{B}_n}(S_1)}{a_{\mathcal{B}_n}(S_1)} = \frac{c_{\mathcal{F}_n} \cdot \Gamma(k_1+1/2)}{\Gamma(k_1)} \quad \text{and} \quad \lim_{n\to\infty} \sqrt{n}\, \frac{g_{\mathcal{B}_n}(S_2)}{a_{\mathcal{B}_n}(S_2)} = \frac{c_{\mathcal{F}_n} \cdot \Gamma(k_2+1/2)}{\Gamma(k_2)},$$

where

$$c_{\mathcal{B}_n} = \sqrt{2}, \qquad c_{\mathcal{V}_n} = b\rho \qquad \text{and} \qquad c_{\mathcal{T}_n} = 2.$$

Note that the values k_1, $k_1 + 1/2$, k_2 and $k_2 + 1/2$ all belong to the set $\{\frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots\}$. First, we are going to show that the function $f(k) = \frac{\Gamma(k+1/2)}{\Gamma(k)}$ is increasing in k for $k \in \{\frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots\}$. Indeed, applying twice Gautschi's inequality (Lemma 3) we get for $k > 1/2$

$$\frac{f(k+1/2)}{f(k)} = \frac{\Gamma(k+1)}{\Gamma(k+1/2)} \frac{\Gamma(k)}{\Gamma(k+1/2)} > k^{1/2}(k+1/2)^{1/2}.$$

Thus, for $k > \frac{\sqrt{17}-1}{4} \approx 0.78$, we obtain $\frac{f(k+1/2)}{f(k)} > 1$. For $k = 1/2$ we also have $\frac{f(k+1/2)}{f(k)} = \frac{\pi}{2} > 1$.

Now, it suffices to show that whenever $S_1 \subseteq S_2$, then $k_1 \leqslant k_2$ (equivalently $m_1 + l_1 \leqslant m_2 + l_2$). Of course, $m_1 \leqslant m_2$. Next, observe that if $S_1 \subseteq S_2$, then also $l_1 \leqslant l_2$. Indeed, the number of leaves in a tree is the cardinality of its largest antichain. If S_1 has l_1 leaves and $S_1 \subseteq S_2$, then S_2 needs to contain an antichain of cardinality l_1 as a subposet, which means that its number of leaves has to satisfy $l_2 \geqslant l_1$. Together we get $m_1 + l_1 \leqslant m_2 + l_2$. □

Theorem 22. *Let S_1, S_2 be rooted trees such that $S_1 \subseteq S_2$ and let $\mathcal{F}_n$ denote any of the families $\mathcal{B}_n$, $\mathcal{V}_n$, $\mathcal{T}_n$. Then for sufficiently large n*

$$\frac{g_{\mathcal{F}_n}(S_1)}{a_{\mathcal{F}_n}(S_1)} \leqslant \frac{g_{\mathcal{F}_n}(S_2)}{a_{\mathcal{F}_n}(S_2)}.$$

Proof. Let $d_{S_1} = (l_1, u_1, \ldots)$, $d_{S_2} = (l_2, u_2, \ldots)$, $k_1 = \frac{m_1+l_1-2}{2}$, $k_2 = \frac{m_2+l_2-2}{2}$. Aiming for a contradiction, assume that $S_1 \subseteq S_2$ and that there is an increasing sequence $n_0 < n_1 < n_2 < \cdots$ such that $\frac{g_{\mathcal{F}_n}(S_1)}{a_{\mathcal{F}_n}(S_1)} > \frac{g_{\mathcal{F}_n}(S_2)}{a_{\mathcal{F}_n}(S_2)}$ for all $n \in \{n_i \mid i \in \mathbb{N}\}$. Then by Theorem 21

$$\lim_{n\to\infty} \sqrt{n}\, \frac{g_{\mathcal{F}_n}(S_1)}{a_{\mathcal{F}_n}(S_1)} = \lim_{n\to\infty} \sqrt{n}\, \frac{g_{\mathcal{F}_n}(S_2)}{a_{\mathcal{F}_n}(S_2)} = \frac{c_{\mathcal{F}_n} \cdot \Gamma(k_1 + 1/2)}{\Gamma(k_1)} = \frac{c_{\mathcal{F}_n} \cdot \Gamma(k_2 + 1/2)}{\Gamma(k_2)},$$

where

$$c_{\mathcal{B}_n} = \sqrt{2}, \qquad c_{\mathcal{V}_n} = b\rho \qquad \text{and} \qquad c_{\mathcal{T}_n} = 2.$$

Recall that the function $f(k) = \frac{\Gamma(k+1/2)}{\Gamma(k)}$ is increasing in k for $k \in \{\frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots\}$ thus the above equality implies $k_1 = k_2$, or equivalently $m_1 + l_1 = m_2 + l_2$. By $S_1 \subseteq S_2$ we have $l_1 \leqslant l_2$ and $m_1 \leqslant m_2$ (see the proof of Theorem 21), therefore we get $l_1 = l_2$ and $m_1 = m_2$. Thus S_1 and S_2 are isomorphic and $\frac{g_{\mathcal{F}_n}(S_1)}{a_{\mathcal{F}_n}(S_1)} = \frac{g_{\mathcal{F}_n}(S_2)}{a_{\mathcal{F}_n}(S_2)}$ which is a contradiction. □

9 Discussion

We proved that the ratio of the number of good embeddings to the number of all embeddings of a given tree $S = (\bullet, S_1, \ldots, S_k)$ into the families of trees $\mathcal{B}_n, \mathcal{V}_n, \mathcal{T}_n$ is asymptotically of the same order for all the three considered families of trees, namely plane binary trees, non-plane binary trees and planted plane trees. Thereby we extended the results of Kubicki *et al.* [29, 30] and Georgiou [22]. We expect that this result will also hold for the family of Pólya trees, which are the closest counterpart to posets that admit a (rooted) tree-like shape, *i.e.* they have a single maximal element and each interval between two of their elements is a chain. In principle, the approach that we used within this paper works for embeddings into the family of Pólya trees as well. However, one would have to consider all possible partitions of $S_1, \ldots, S_k$, as any collection of isomorphic subtrees within $S_1, \ldots, S_k$ admits non-trivial isomorphisms between the S_i's, which can get rather involved and is therefore omitted in this work.

Acknowledgements

The authors thank an anonymous referee for pointing out a subtle error in the first version of the paper and a second referee for pointing out further references and for numerous comments on the presentation that led to a substantial improvement of the paper.

References

[1] Noga Alon and Joel H. Spencer. *The Probabilistic Method.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2008.

[2] Andrei Asinowski, Axel Bacher, Cyril Banderier, and Bernhard Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns, the vectorial kernel method, and generating functions for pushdown automata. *Algorithmica*, 82(3):386–428, 2020.

[3] Fabrício Siqueira Benevides and Małgorzata Sulkowska. Percolation and best-choice problem for powers of paths. *J. Applied Probability*, 54(2):343–362, 2017.

[4] Fabrício Siqueira Benevides and Malgorzata Sulkowska. Maximizing the expected number of components in an online search of a graph. *Discrete Math.*, 345(1):112668, 2022.

[5] Miklós Bóna. *Combinatorics of Permutations.* Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, second edition, 2012.

[6] Frédéric Chyzak, Michael Drmota, Thomas Klausner, and Gerard Kok. The distribution of patterns in random trees. *Combin. Probab. Comput.*, 17(1):21–59, 2008.

[7] Gwendal Collet, Michael Drmota, and Lukas D. Klausner. Limit laws of planar maps with prescribed vertex degrees. *Combin. Probab. Comput.*, 28(4):519–541, 2019.

[8] Michael Dairyko, Lara Pudwell, Samantha Tyner, and Casey Wynn. Non-contiguous pattern avoidance in binary trees. *Electron. J. Combin.*, 19(3), #P22, 2012.

[9] Emeric Deutsch. Dyck path enumeration. *Discrete Math.*, 204(1-3):167–202, 1999.

[10] Michael Drmota. *Random Trees.* Springer, Vienna-New York, 2009.

[11] Michael Drmota and Bernhard Gittenberger. The distribution of nodes of given degree in random trees. *J. Graph Theory*, 31(3):227–253, 1999.

[12] Michael Drmota, Lander Ramos, and Juanjo Rué. Subgraph statistics in subcritical graph classes. *Random Struct. Algor.*, 51(4):631–673, 2017.

[13] Michael Drmota and Benedikt Stufler. Pattern occurrences in random planar maps. *Statist. Probab. Lett.*, 158, 2020.

[14] Michael Drmota and Guan-Ru Yu. The number of double triangles in random planar maps. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 110 of *LIPIcs. Leibniz Int. Proc. Inform.* Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.

[15] Thomas S. Ferguson. Who solved the secretary problem? *Statist. Sci.*, 4(3):282–289, 1989.

[16] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.

[17] Peter R. Freeman. The secretary problem and its extensions: A review. *Int. Stat. Rev.*, 51:189–206, 1983.

[18] Ragnar Freij and Johan Wästlund. Partially ordered secretaries. *Electron. Commun. Probab.*, 15:504–507, 2010.

[19] Anton Freund. From Kruskal's theorem to Friedman's gap condition. *Math. Struct. Comp. Sci.*, 30(8):952–975, 2020.

[20] Bryn Garrod and Robert Morris. The secretary problem on an unknown poset. *Random Struct. Algor.*, 43(4):429–451, 2013.

[21] Walter Gautschi. Some elementary inequalities relating to the gamma and incomplete gamma function. *J. Math. Phys. Camb.*, 38:77–81, 1959.

[22] Nicholas Georgiou. Embeddings and other mappings of rooted trees into complete trees. *Order*, 22(3):257–288, 2005.

[23] John P. Gilbert and Frederick Mosteller. Recognizing the maximum of a sequence. In *Selected Papers of Frederick Mosteller*, pages 355–398. Springer, 2006.

[24] Bernhard Gittenberger. Nodes of large degree in random trees and forests. *Random Struct. Algor.*, 28(3):374–385, 2006.

[25] Alexander V. Gnedin. Multicriteria extensions of the best choice problem: Sequential selection without linear order. *Contemp. Math.*, 125:153–172, 1992.

[26] Andrzej Grzesik, Michał Morayne, and Małgorzata Sulkowska. From directed path to linear order—the best choice problem for powers of directed paths. *SIAM J. Discrete Math.*, 29(1):500–513, 2015.

[27] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random Graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[28] Wojciech Kaźmierczak. The best choice problem for a union of two linear orders with common maximum. *Discrete Appl. Math.*, 161(18):3090–3096, 2013.

[29] Grzegorz Kubicki, Jenö Lehel, and Michał Morayne. A ratio inequality for binary trees and the best secretary. *Combin. Probab. Comput.*, 11(2):149–161, 2002.

[30] Grzegorz Kubicki, Jenö Lehel, and Michał Morayne. An asymptotic ratio in the complete binary tree. *Order*, 20(2):91–97, 2003.

[31] Grzegorz Kubicki, Jeno Lehel, and Michał Morayne. Counting chains and antichains in the complete binary tree. *Ars Combinatoria*, 79, 2006.

[32] Grzegorz Kubicki and Michał Morayne. Graph-theoretic generalization of the secretary problem: The directed path case. *SIAM J. Discrete Math.*, 19(3):622–632, 2005.

[33] Małgorzata Kuchta, Michał Morayne, and Jarosław Niemiec. Counting embeddings of a chain into a tree. *Discrete Math.*, 297(1-3):49–59, 2005.

[34] Małgorzata Kuchta, Michał Morayne, and Jarosław Niemiec. Counting embeddings of a chain into a binary tree. *Ars Combinatoria*, 91, 2009.

[35] Xueliang Li, Yiyang Li, and Yongtang Shi. The asymptotic number of non-isomorphic rooted trees obtained by rooting a tree. *J. Math. Anal. Appl.*, 434(1):1–11, 2016.

[36] Dennis V. Lindley. Dynamic programming and decision theory. *Appl. Stat. - J. Roy. St. C*, 10(1):39–51, 1961.

[37] Michał Morayne. Partial-order analogue of the secretary problem - the binary tree case. *Discrete Math.*, 184(1-3):165–181, 1998.

[38] Marc Noy, Clément Requilé, and Juanjo Rué. Further results on random cubic planar graphs. *Random Struct. Algor.*, 56(3):892–924, 2020.

[39] Richard Otter. The number of trees. *Ann. Math.*, 49(2):583–599, 1948.

[40] Konstantinos Panagiotou and Makrand Sinha. Vertices of degree k in random unlabeled trees. *J. Graph Theory*, 69(2):114–130, 2012.

[41] John Preater. The best-choice problem for partially ordered objects. *Oper. Res. Lett.*, 25(4):187–190, 1999.

[42] Robert W. Robinson and Allen J. Schwenk. The distribution of degrees in a large random tree. *Discrete Math.*, 12(4):359–372, 1975.

[43] Eric S. Rowland. Pattern avoidance in binary trees. *J. Combin. Theory Ser. A*, 117(6):741–758, 2010.

[44] Stephen M. Samuels. Secretary problems. In *Handbook of Sequential Analysis*, volume 118 of *Statist. Textbooks Monogr.*, pages 381–405. Dekker, New York, 1991.

[45] Wolfgang Stadje. Efficient stopping of a random series of partially ordered points. *Multiple Criteria Decision Making Theory and Application. Lect. Notes Econ. Math.*, 177:430–447, 1980.

[46] Benedikt Stufler. Local convergence of random planar graphs. *J. Eur. Math. Soc.*, 2021.

[47] Małgorzata Sulkowska. The best choice problem for upward directed graphs. *Discrete Optim.*, 9(3):200–204, 2012.

[48] William T. Trotter. Partially ordered sets. In *Handbook of Combinatorics, Vol. 1, 2*, pages 433–480. Elsevier Sci. B. V., Amsterdam, 1995.