CONGRUENCE PROPERTIES OF DEPTHS IN SOME RANDOM TREES

SVANTE JANSON

Abstract. Consider a random recursive tree with n vertices. We show that the number of vertices with even depth is asymptotically normal as $n \to \infty$. The same is true for the number of vertices of depth divisible by m for $m = 3$, 4 or 5; in all four cases the variance grows linearly. On the other hand, for $m \geq 7$, the number is not asymptotically normal, and the variance grows faster than linear in n. The case $m = 6$ is intermediate: the number is asymptotically normal but the variance is of order $n \log n$.

This is a simple and striking example of a type of phase transition that has been observed by other authors in several cases. We prove, and explain, this non-intuitive behaviour using a translation to a generalized Pólya urn.

Similar results hold for a random binary search tree; now the number of vertices of depth divisible by m is asymptotically normal for $m \leq 8$ but not for $m \geq 9$, and the variance grows linearly in the first case both faster in the second. (There is no intermediate case.)

In contrast, we show that for conditioned Galton–Watson trees, including random labelled trees and random binary trees, there is no such phase transition: the number is asymptotically normal for every m.

1. Introduction

Given a rooted tree T, let $X_j(T)$ be the number of vertices of depth j (i.e., of distance j from the root). The sequence $(X_j)_{j=0}^{\infty}$ is called the *profile* of the tree and has been studied for various types of random trees by many authors, see e.g. [1, 7, 8, 14, 15, 16, 21, 30].

We will here study the congruence class of the depth modulo some given integer $m \geq 2$. Thus, let $X_j^{(m)}(T) := \sum_{k \equiv j \pmod{m}} X_k(T)$ be the number of vertices of depth congruent to j modulo m. For example, $X_0^{(2)}(T)$ is the number of vertices of even depth. We further let $\mathbf{X}^{(m)}(T)$ denote the vector $\big(X_0^{(m)}(T), \dots, X_{m-1}^{(m)}(T)\big)$.

The purpose of this paper is to study the asymptotic distribution as $n \to \infty$ of the random vector $\mathbf{X}^{(m)}(T_n)$ for some random trees T_n with n vertices. For a random recursive tree (RRT), see Section 2 for definitions,

Date: September 21, 2005; revised June 15, 2006.

2000 *Mathematics Subject Classification.* 60C05; 05C05.

Key words and phrases. random recursive tree, random, binary search tree, conditioned Galton–Watson tree, generalized Polya urn, depths of vertices.

we will show (Theorem 2.1) that $\mathbf{X}^{(m)}$ is asymptotically normal for $m \leq 6$ but not for $m \geq 7$; furthermore, the variance grows linearly for $m \leq 5$ but faster for $m \geq 6$. For a random binary search tree (BST), the result is similar (Theorem 2.7): $\mathbf{X}^{(m)}$ is asymptotically normal for $m \leq 8$, but not for $m \geq 9$; moreover, the variance grows linearly for $m \leq 8$ but faster for larger m. In contrast, for a conditioned Galton–Watson tree (CGWT), $\mathbf{X}^{(m)}$ is asymptotically normal for every m (Theorem 2.12).

Note that the typical depths are of the order $\log n$ in a RRT or BST, but $\sqrt{n}$ in a CGWT. There is thus more room for smoothing between the congruence classes in the latter case, which may explain the asymptotic normality in that case, but we see no intuitive explanation for the difference between small and large m for the RRT and BST.

We do not claim that the variables $X_j^{(m)}$ have any importance in applications, but they provide a simple and surprising example of a type of phase transition that has been observed in several similar combinatorial situations. One well-known such example is random m-ary search trees, see Chern and Hwang [11]; see also e.g. Mahmoud and Pittel [31], Lew and Mahmoud [29], Fill and Kapur [17], Chauvin and Pouyanne [9]. Other examples are random quadtrees, see Chern, Fuchs and Hwang [10], and random fragmentation trees, see Dean and Majumdar [12] and Janson and Neininger [26].

We will use a translation into generalized Pólya urns; for such urns, there is a general theorem describing the phase transition in terms of the size of the real part of the second largest eigenvalue, see Athreya and Karlin [2] or Athreya and Ney [3, §V.9]; see also Kesten and Stigum [27] and, for more details, [25]. The same type of phase transition in a related (and overlapping) setting is described in [23]. We hope that our simple example can help illustrate this surprising and non-intuitive phenomenon. It provides also an illustration of the results in [25] in a simple concrete situation.

We state the results in Section 2. Proofs are given in the following three sections, with one type of random trees in each. The proofs for RRT and BST in Sections 3 and 4 are very similar and uses generalized Pólya urns, while the proof for CGWT in Section 5 uses generating functions and singularity analysis. Finally, in Section 6 we show that the oscillations shown in the results for RRT and BST except for small m are genuine and not only an artefact of the proof or of the normalization.

Acknowledgement. This research was mainly done during the workshop *Probability theory on trees and analysis of algorithms* at the Mathematisches Forschungsinstitut Oberwolfach in August 2004. I thank the participants, in particular Hsien-Kuei Hwang, for inspiring comments.

2. Definitions and results

Let $\mathbf{1}$ denote the m-dimensional vector $(1, \dots, 1)$. (We do not always distinguish between row and column vectors in our notation.)

We will occasionally deal with complex random variables. A *complex Gaussian variable* is a complex random variable ζ such that $\operatorname{Re}\zeta$ and $\operatorname{Im}\zeta$ are jointly Gaussian (i.e., normal). A complex Gaussian variable ζ is *symmetric* if and only if $\mathbb{E}\,\zeta = \mathbb{E}\,\zeta^2 = 0$; its distribution is then determined by the scale factor $\mathbb{E}\,|\zeta|^2$. See further [24, Section I.4].

The *discrete Fourier transform* on the group $\mathbb{Z}_m = \{0, \dots, m-1\}$ is defined by

$$\widehat{f}(k) = \sum_{j=0}^{m-1} \omega^{kj} f(j), \tag{2.1}$$

where $\omega = \omega_m := e^{2\pi \mathrm{i}/m}$.

All limits below are as $n \to \infty$.

2.1. **Random recursive trees.** A *random recursive tree (RRT)* with n vertices is a random rooted tree obtained by starting with a single root and then adding $n-1$ vertices one by one, each time joining the new vertex to a randomly chosen old vertex; the random choices are uniform and independent of each other. If the vertices are labelled $1, 2, \dots$, we then obtain a tree where the labels increase along each branch as we travel from the root; the random recursive tree can also be defined as a (uniform) randomly chosen such labelled tree. (The distribution of a random recursive tree differs from the distribution of a uniform random labelled tree.) See also the survey [33].

We state the main results for RRT in the following theorem, and add various details in the remarks after it.

Theorem 2.1. *Let T_n be a random recursive tree with n vertices. Then, the following holds for $\mathbf{X}^{(m)} = \mathbf{X}^{(m)}(T_n)$, as $n \to \infty$.*

(i) *If $2 \le m \le 5$, then*

$$n^{-1/2}\Bigl(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\Bigr) \xrightarrow{\mathrm{d}} N(0, \Sigma_m) \tag{2.2}$$

for a covariance matrix Σ_m given explicitly in (2.5)–(2.8)*.*

(ii) *If $m = 6$, then*

$$(n \ln n)^{-1/2}\Bigl(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\Bigr) \xrightarrow{\mathrm{d}} N(0, \Sigma_m),$$

where Σ_m is given explicitly in (2.9)*.*

(iii) *For $m \ge 7$, let $\alpha = \cos(2\pi/m) > 1/2$ and $\beta = \sin(2\pi/m)$. Then*

$$n^{-\alpha}\Bigl(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\Bigr) - \operatorname{Re}\bigl(n^{\mathrm{i}\beta}\widetilde{W}_m\bigr) \xrightarrow{\mathrm{a.s.}} 0, \tag{2.3}$$

for some complex random vector $\widetilde{W}_m = \bigl(\widetilde{Z}_m e^{-2\pi j \mathrm{i}/m}\bigr)_{j=0}^{m-1}$, where $\widetilde{Z}_m$ is a complex random variable. In particular, along any subsequence such that $\beta \ln n \mod 2\pi \to \gamma$ for some $\gamma \in [0, 2\pi]$,

$$n^{-\alpha}\Bigl(X_j^{(m)} - \frac{n}{m}\Bigr) \xrightarrow{\mathrm{d}} \operatorname{Re}\bigl(e^{\mathrm{i}(\gamma - 2\pi j/m)}\widetilde{Z}_m\bigr), \tag{2.4}$$

jointly in $j = 0, \dots, m-1$.

Remark 2.2. The covariance matrices Σ_m in (i) and (ii) are given by

$$\Sigma_2 = \frac{1}{12}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{2.5}$$

$$\Sigma_3 = \frac{1}{18}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \tag{2.6}$$

$$\Sigma_4 = \frac{1}{48}\begin{pmatrix} 7 & -1 & -5 & -1 \\ -1 & 7 & -1 & -5 \\ -5 & -1 & 7 & -1 \\ -1 & -5 & -1 & 7 \end{pmatrix} \tag{2.7}$$

$$\Sigma_5 = \frac{1}{25}\begin{pmatrix} 6 & 1 & -4 & -4 & 1 \\ 1 & 6 & 1 & -4 & -4 \\ -4 & 1 & 6 & 1 & -4 \\ -4 & -4 & 1 & 6 & 1 \\ 1 & -4 & -4 & 1 & 6 \end{pmatrix} \tag{2.8}$$

$$\Sigma_6 = \frac{1}{36}\begin{pmatrix} 2 & 1 & -1 & -2 & -1 & 1 \\ 1 & 2 & 1 & -1 & -2 & -1 \\ -1 & 1 & 2 & 1 & -1 & -2 \\ -2 & -1 & 1 & 2 & 1 & -1 \\ -1 & -2 & -1 & 1 & 2 & 1 \\ 1 & -1 & -2 & -1 & 1 & 2 \end{pmatrix} \tag{2.9}$$

Note that the matrices Σ_m are circulant, i.e. invariant under a cyclic shift of both rows and columns. In particular, $X_j^{(m)}$ has the same asymptotic distribution for every j (when $m \le 6$).

Remark 2.3. The distribution of $\widetilde{Z}_m = \widetilde{Z}$ in (iii) is determined by the equalities in distribution

$$Z \overset{\mathrm{d}}{=} \tfrac{m}{2} W^{\omega} \widetilde{Z}, \tag{2.10}$$

$$Z \overset{\mathrm{d}}{=} U^{\omega}(Z + \omega Z'), \tag{2.11}$$

where $\omega = \alpha + \beta \mathrm{i} = e^{2\pi \mathrm{i}/m}$, $Z' \overset{\mathrm{d}}{=} Z$, $W \sim \operatorname{Exp}(1)$, $U \sim U(0,1)$, and W, $\widetilde{Z}$, U, Z, Z' are independent, together with $\mathbb{E}\, Z = 1$.

The moments of Z and $\widetilde{Z}$ are all finite, and can be obtained recursively from (2.11) and (2.10) together with $\mathbb{E}\, Z = 1$. In particular,

$$\begin{aligned}
\mathbb{E}\, Z &= 1, & \mathbb{E}\, \widetilde{Z} &= \frac{2}{m\Gamma(1+\omega)}, \\
\mathbb{E}\, Z^2 &= \frac{2}{2-\omega}, & \mathbb{E}\, \widetilde{Z}^2 &= \frac{8}{m^2(2-\omega)\Gamma(1+2\omega)}, \\
\mathbb{E}\, |Z|^2 &= \frac{2\alpha}{2\alpha-1}, & \mathbb{E}\, |\widetilde{Z}|^2 &= \frac{8\alpha}{m^2(2\alpha-1)\Gamma(1+2\alpha)}, \\
\mathbb{E}\, Z^3 &= \frac{6(1+\omega)}{(3-\omega^2)(2-\omega)}, & \mathbb{E}\, \widetilde{Z}^3 &= \frac{48(1+\omega)}{m^3(3-\omega^2)(2-\omega)\Gamma(1+3\omega)}.
\end{aligned}$$

Remark 2.4. It follows from [23, Section 2] that all moments of the variables in Theorem 2.1 stay bounded as $n \to \infty$, and thus moment convergence holds in (i) and (ii) and for convergent subsequences in (iii). In particular, the variance of each $X_j^{(m)}$ is of order n for $m \le 5$, but larger for $m \ge 6$.

Remark 2.5. The covariance matrix Σ_m has rank $m-1$ when $m \le 5$, while Σ_6 has rank 2 only. (Full rank is impossible because all row sums are 0, reflecting the fact that the total number of vertices is non-random.)

Remark 2.6. The results for the RRT T_n become simpler if we state them in terms of the discrete Fourier transform $\widehat{\mathbf{X}}^{(m)} = \big(\widehat{X}^{(m)}(k)\big)_0^{m-1}$ defined by (2.1). (See also [23], where the proof is based on this Fourier transform.) Clearly, $\widehat{X}^{(m)}(0) = \sum_j X_j^{(m)} = |T_n| = n$ is deterministic.

For $2 \le m \le 5$, Theorem 2.1 implies (and is equivalent to) the joint convergence

$$n^{-1/2}\widehat{X}^{(m)}(k) \overset{\mathrm{d}}{\longrightarrow} V_k, \qquad k = 1, \dots, m-1, \tag{2.12}$$

where V_k are complex (jointly) Gaussian variables such that, for $k, l \in \{1, \dots, m-1\}$, using (3.1) below,

(i) $V_{m-k} = \overline{V}_k$,
(ii) $\mathbb{E}\, V_k = 0$,
(iii) $\mathbb{E}(V_k V_l) = 0$ when $l \ne m-k$,
(iv) $\mathbb{E}\, |V_k|^2 = \dfrac{1}{1 - 2\operatorname{Re}\omega_m^k} = \dfrac{1}{1-2\cos(2\pi k/m)}$.

It follows that further:

(v) If $k = m/2$ (and thus $\omega_m^k = -1$), then $V_k = \overline{V}_k$ is a symmetric real Gaussian variable.
(vi) If $k \ne m/2$, then $\mathbb{E}\, V_k^2 = 0$ and thus V_k is a symmetric complex Gaussian variable.
(vii) The Gaussian variables V_k, $1 \le k \le m/2$, are independent.

Note that the joint distribution of $V_1, \dots, V_{m-1}$ is determined by (iv)–(vii) together with (i).

Similarly, for $m = 6$,

$$(n \ln n)^{-1/2} \widehat{X}^{(m)}(k) \xrightarrow{\text{d}} V_k, \qquad k = 1, \dots, 5,$$

where now, however, only V_1 and $V_5 = \overline{V_1}$ are non-zero; V_1 is a symmetric complex Gaussian variable with $\mathbb{E}\,|V_1|^2 = 1$.

Finally, for $m \geq 7$, we have

$$n^{-\omega_m} \widehat{X}^{(m)}(1) \xrightarrow{\text{a.s.}} \frac{m}{2} \widetilde{Z}_m, \tag{2.13}$$

while $n^{-\omega_m} \widehat{X}^{(m)}(k) \xrightarrow{\text{a.s.}} 0$ for $k = 2, \dots, m-2$.

2.2. **Binary search trees.** A *binary search tree (BST)* is constructed from a sequence $x_1, \dots, x_n$ of distinct real numbers as follows, see e.g. [28, Section 6.2.2]. If $n = 0$, the tree is empty. Otherwise, start with a root. (In computer applications, x_1 is stored in the root.) Then construct recursively two subtrees of the root by the same procedure applied to two subsequences of $x_1, \dots, x_n$: the left subtree from the x_i with $x_i < x_1$ and the right subtree from the x_i with $x_i > x_1$. The number of vertices in the tree is thus n, and each vertex corresponds to an x_i.

A random BST is obtained by this construction applied to a sequence $x_1, \dots, x_n$ in random order. (Since only the order properties of $x_1, \dots, x_n$ matter, we can let them be, for example, either a random permutation of $1, \dots, n$ or n i.i.d. random variables with a common continuous distribution.)

It is easily seen that a random BST can be grown by adding vertices one by one according to a Markov process similar to the definition of the RRT: Given a binary tree with n vertices, there are $n+1$ possible positions for a new vertex, and we choose one of them at random (uniformly).

Theorem 2.7. *Let T_n be a random binary search tree with n vertices. Then, the following holds for $\mathbf{X}^{(m)} = \mathbf{X}^{(m)}(T_n)$, as $n \to \infty$.*

(i) *If $2 \leq m \leq 8$, then*

$$n^{-1/2}\Bigl(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\Bigr) \xrightarrow{\text{d}} N(0, \Sigma_m)$$

for a covariance matrix Σ_m given explicitly in Remark 2.8.

(ii) *For $m \geq 9$, let $\alpha = 2\cos(2\pi/m) - 1 > 1/2$ and $\beta = 2\sin(2\pi/m)$. Then*

$$n^{-\alpha}\Bigl(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\Bigr) - \operatorname{Re}\bigl(n^{\mathrm{i}\beta}\widetilde{W}_m\bigr) \xrightarrow{\text{a.s.}} 0, \tag{2.14}$$

for some complex random vector $\widetilde{W}_m = \bigl(\widetilde{Z}_m e^{-2\pi j \mathrm{i}/m}\bigr)_{j=0}^{m-1}$, where $\widetilde{Z}_m$ is a complex random variable. In particular, along any subsequence such that $\beta \ln n \mod 2\pi \to \gamma$ for some $\gamma \in [0, 2\pi]$,

$$n^{-\alpha}\Bigl(X_j^{(m)} - \frac{n}{m}\Bigr) \xrightarrow{\text{d}} \operatorname{Re}\bigl(e^{\mathrm{i}(\gamma - 2\pi j/m)}\widetilde{Z}_m\bigr), \tag{2.15}$$

jointly in $j = 0, \dots, m-1$.

Remark 2.8. The covariance matrices Σ_m in (i) are circulant and explicitly given by the following first rows:

$$
\begin{aligned}
m=2:&\quad \tfrac{1}{28}(1,-1),\\
m=3:&\quad \tfrac{1}{45}(2,-1,-1),\\
m=4:&\quad \tfrac{1}{336}(17,-3,-11,-3),\\
m=5:&\quad \tfrac{1}{275}(16,1,-9,-9,1),\\
m=6:&\quad \tfrac{1}{1260}(89,23,-37,-61,-37,23),\\
m=7:&\quad \tfrac{1}{637}(62,27,-15,-43,-43,-15,27),\\
m=8:&\quad \tfrac{1}{1344}(269,165,-11,-171,-235,-171,-11,165).
\end{aligned}
$$

Remark 2.9. The distribution of $\widetilde{Z}_m = \widetilde{Z}$ in (ii) is determined by the equalities in distribution

$$Z \overset{\mathrm{d}}{=} \tfrac{m}{2}(2\omega-1)W^{2\omega-1}\widetilde{Z}, \tag{2.16}$$

$$Z \overset{\mathrm{d}}{=} \omega U^{2\omega-1}(Z+Z'), \tag{2.17}$$

where $\omega=\omega_m=e^{2\pi\mathrm{i}/m}$, $Z' \overset{\mathrm{d}}{=} Z$, $W\sim \mathrm{Exp}(1)$, $U\sim U(0,1)$, and W, $\widetilde{Z}$, U, Z, Z' are independent, together with $\mathbb{E}\,Z=1$.

Again, the moments of Z and $\widetilde{Z}$ are all finite, and can be obtained recursively from (2.17) and (2.16) together with $\mathbb{E}\,Z=1$. In particular,

$$
\begin{aligned}
\mathbb{E}\,Z &= 1, & \mathbb{E}\,\widetilde{Z} &= \frac{2}{m(2\omega-1)\Gamma(2\omega)},\\
\mathbb{E}\,Z^2 &= \frac{2\omega^2}{4\omega-1-2\omega^2}, & \mathbb{E}\,\widetilde{Z}^2 &= \frac{4}{m^2(2\omega-1)^2\Gamma(4\omega-1)}\,\mathbb{E}\,Z^2,\\
\mathbb{E}\,|Z|^2 &= \frac{2}{4\alpha-3}, & \mathbb{E}\,|\widetilde{Z}|^2 &= \frac{4}{m^2(5-4\alpha)\Gamma(4\alpha-1)}\,\mathbb{E}\,|Z|^2,\\
\mathbb{E}\,Z^3 &= \frac{6\omega^5}{(3\omega-1-\omega^3)(4\omega-1-\omega^2)}, & \mathbb{E}\,\widetilde{Z}^3 &= \frac{8}{m^3(2\omega-1)^3\Gamma(6\omega-2)}\,\mathbb{E}\,Z^3.
\end{aligned}
$$

Remark 2.10. The proofs in [23, Section 2] apply to this situation too and show that all moments of the variables in Theorem 2.7 stay bounded as $n\to\infty$, and thus moment convergence holds in (i) and for convergent subsequences in (ii). In particular, the variance of each $X_j^{(m)}$ is of order n for $m\le 8$, but larger for $m\ge 9$.

Remark 2.11. Just as for the RRT case in Remark 2.6, the results become simpler if we state them in terms of the Fourier transform $\widehat{\mathbf{X}}^{(m)}$. Clearly, again $\widehat{X}^{(m)}(0)=|T_n|=n$.

Further, for $2\le m\le 8$, (2.12) holds, where V_k are complex Gaussian variables satisfying (i)–(iii) and (v)–(vii) in Remark 2.6 together with

(iv′) $\mathbb{E}\,|V_k|^2 = \dfrac{1}{3-4\cos(2\pi k/m)}$.

Finally, for $m \geq 9$,

$$n^{-(2\omega_m-1)}\widehat{X}^{(m)}(1) \xrightarrow{\text{a.s.}} \frac{m}{2}\widetilde{Z}_m, \tag{2.18}$$

while $n^{-(2\omega_m-1)}\widehat{X}^{(m)}(k) \xrightarrow{\text{a.s.}} 0$ for $k = 2, \dots, m-2$.

2.3. **Conditioned Galton–Watson trees.** A *conditioned Galton–Watson tree (CGWT)* with n vertices is a random tree obtained as the family tree of a Galton–Watson process conditioned on a given total population of n. (See e.g. [1, 13] for details.) The Galton–Watson process is defined using an offspring distribution; let ξ denote a random variable with this distribution. We assume, as usual, $\mathbb{E}\,\xi = 1$ (the Galton–Watson process is critical) and $0 < \sigma^2 = \operatorname{Var}\xi < \infty$. We assume further, for technical reasons, that $\mathbb{E}\,e^{a\xi} < \infty$ for some $a > 0$. (Equivalently, the probability generating function $\varphi(z) := \mathbb{E}\,z^{\xi}$ is analytic in a disc with radius greater than 1.)

It is well-known [1] that the conditioned Galton–Watson trees are the same as the simply generated trees [32]. Many combinatorially interesting random trees are of this type, with different choices of ξ, for example labelled trees ($\xi \sim \mathrm{Po}(1)$, $\sigma^2 = 1$); ordered (=plane) trees ($\mathbb{P}(\xi = k) = 2^{-k-1}$, $\sigma^2 = 2$); binary trees ($\xi \sim \mathrm{Bi}(2, 1/2)$, $\sigma^2 = 1/2$); complete binary trees ($\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$, $\sigma^2 = 1$).

It has been shown by Drmota and Gittenberger [14] that the profile of a CGWT converges, after normalization, to the local time of a Brownian excursion. For the congruence classes we have a simpler result. As in many other results, see e.g. [1, 14], different choices of ξ only affects a scaling factor in the limit result.

Theorem 2.12. *Let T_n be a random conditioned Galton–Watson tree with n vertices, with the offspring distribution given by a random variable ξ such that $\mathbb{E}\,\xi = 1$, $\sigma^2 := \operatorname{Var}\xi > 0$ and $\mathbb{E}\,e^{a\xi} < \infty$ for some $a > 0$. Then, for any fixed $m \geq 2$, the following holds for $\mathbf{X}^{(m)} = \mathbf{X}^{(m)}(T_n)$ as $n \to \infty$,*

$$n^{-1/2}\Bigl(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\Bigr) \xrightarrow{\text{d}} N(0, \Sigma_m)$$

for a covariance matrix $\Sigma_m = (\sigma_{ij})$ given explicitly by

$$\sigma_{i,i+k} = \frac{\sigma^2}{12m^2}\bigl(m^2 - 1 - 6k(m-k)\bigr)$$

for $0 \leq i \leq m-1$, $0 \leq k \leq m$ and $i + k$ taken modulo m.

Remark 2.13. The covariance matrices Σ_m are circulant. For small m, they are explicitly given by the following first rows:

$$\begin{aligned} m=2: &\quad \tfrac{\sigma^2}{16}(1,-1), \\ m=3: &\quad \tfrac{\sigma^2}{27}(2,-1,-1), \\ m=4: &\quad \tfrac{\sigma^2}{64}(5,-1,-3,-1), \\ m=5: &\quad \tfrac{\sigma^2}{25}(2,0,-1,-1,0), \\ m=6: &\quad \tfrac{\sigma^2}{432}(35,5,-13,-19,-13,5). \end{aligned}$$

Remark 2.14. It follows from the proof in Section 5 that moment convergence holds in Theorem 2.12. In particular, the variance of each $X_j^{(m)}$ is of order n for every m.

Remark 2.15. Note the curious fact that the asymptotic variance of $X_j^{(m)}$, $n\sigma^2(1-m^{-2})/12$, is almost independent of m.

Remark 2.16. For the Fourier transform $\widehat{\mathbf{X}}^{(m)}$, we again have $\widehat{X}^{(m)}(0) = |T_n| = n$. Further, for any $m \geq 2$, (2.12) holds, where V_k are complex Gaussian variables satisfying (i)–(iii) and (v)–(vii) in Remark 2.6 together with

(iv″) $\mathbb{E}\,|V_k|^2 = \dfrac{1}{|1-\omega_m^k|^2} = \dfrac{1}{2-2\cos(2\pi k/m)}$.

Remark 2.17. As a comparison to the results above, suppose that we instead of considering depths modulo m give each vertex a random label $0,\dots,m-1$. Equivalently, we construct $\mathbf{X}^{(m)} = (X_0^{(m)},\dots,X_{m-1}^{(m)})$ by randomly throwing n balls into m urns $0,\dots,m-1$, letting $X_j^{(m)}$ be the number of balls in urn j. Then $X_j^{(m)} \sim \mathrm{Bi}(n,1/m)$, and the central limit theorem shows that $n^{-1/2}\bigl(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\bigr) \xrightarrow{\mathrm{d}} N(0,\Sigma_m)$ with $\Sigma = (m^{-1}\delta_{ij} - m^{-2})_{ij}$. Moreover, the Fourier transform $\widehat{\mathbf{X}}^{(m)}$ is a sum of n i.i.d. complex random variables and the central limit theorem shows that (2.12) holds, where now V_k are complex Gaussian variables satisfying (i)–(iii) and (v)–(vii) in Remark 2.6 together with

(iv‴) $\mathbb{E}\,|V_k|^2 = 1$.

It follows from the formulas above that the asymptotic variance of $X_j^{(m)}(T_n)$ (for any j) for one of the random trees T_n that we consider above is smaller the asymptotic variance of $X_j^{(m)}$ for a random labelling when m is small, but not when m is large. More precisely, it is smaller for RRT when $m \leq 5$ and for BST when $m \leq 7$; for CGWT, it is smaller when $m < 12/\sigma^2 - 1$. (Thus, if $\sigma^2 > 4$, then the asymptotic variance for the CGWT is always larger than for a random labelling.) For these values of m, and only these, we thus find less randomness in the collection of depths modulo m than for a random labelling.

3. Random recursive trees

The definition of RRT in Section 2 shows immediately that the distribution of depths modulo m is given by the following generalized Pólya urn: The urn contains balls with labels $0, \dots, m-1$, representing the depths modulo m of the vertices. Start with a single ball with label 0 in the urn. Then, repeatedly, draw a ball (at random) from the urn, replace it, and if the drawn ball had label j, add a new ball with label $j+1 \pmod m$.

This urn was studied briefly in [25, Example 7.9], see also [23, Example 6.3]. We repeat the analysis in [25] with more details. Using the notation there we have, with the indices in $\{0, \dots, m-1\}$ and addition taken modulo m,

$$\xi_{ij} = \delta_{i+1,j}, \qquad a_j = 1, \qquad A = (\xi_{ji})_{i,j=0}^{m-1} = (\delta_{i,j+1})_{i,j=0}^{m-1}.$$

(Here, ξ_{ij} is the number of balls of type j added when a ball of type i is drawn; in this case, these are deterministic numbers. The matrix $A = (\xi_{ji})$ describes the evolution.) The matrix A is circulant and corresponds to a convolution operator in $\ell^2(\mathbb{Z}_m)$; hence, A is diagonalized by the characters of the group $\mathbb{Z}_m$. More precisely, let $\omega = \omega_m := e^{2\pi \mathrm{i}/m}$. Then A has left eigenvectors $u_j = (\omega^{jk})_{k=0}^{m-1}$ and right eigenvectors $v_j = (m^{-1}\omega^{-jk})_{k=0}^{m-1}$ with eigenvalues ω^j, for $j = 0, \dots, m-1$. Thus, A has m simple eigenvalues, and the dominant eigenvalue λ_1 is $\omega^0 = 1$. (Note that the corresponding eigenvectors $u_0 = \mathbf{1}$ and $v_0 = m^{-1}\mathbf{1}$ are denoted u_1 and v_1 in [25].) We have chosen the normalizations such that $(u_j)_j$ and $(v_j)_j$ are dual bases; moreover u_0 and v_0 satisfy the normalizations in [25, (2.2)–(2.3)].

The eigenvalues with second largest real part are $\lambda_2 = \omega$ and $\lambda_3 = \overline{\omega}$. Since $\operatorname{Re}\lambda_2 = \cos(2\pi/m) < 1/2$ when $m \le 5$, $\operatorname{Re}\lambda_2 = \cos(2\pi/m) = 1/2$ when $m = 6$, and $\operatorname{Re}\lambda_2 = \cos(2\pi/m) > 1/2$ when $m \ge 7$, the trichotomy in Theorem 2.1 follows from [25, Theorems 3.22–3.24]. (The conditions (A1)–(A6) there are easily verified, see [25, pp. 180–181].)

More precisely, in Case (i), the convergence to a normal distribution follows by [25, Theorems 3.22]. To find the covariance matrices Σ_m, we use [25, Lemma 5.3(iii)] (or [25, Lemma 5.3(i) or (ii), Lemma 5.4 and Lemma 5.5]) . We have $D = m^{-1}I$ and thus $u_j' D u_k = m^{-1} u_j \cdot u_k = 0$ unless $j \equiv -k \pmod m$, while $u_{m-k}' D u_k = u_k^* D u_k = 1$. (We use the notation in [25] that $u \cdot v := u^t v$, without complex conjugation.) Hence,

$$\Sigma_m = \sum_{j=1}^{m-1} \frac{1}{1 - 2\operatorname{Re}\omega_m^j} v_j v_j^*, \tag{3.1}$$

and a straightforward evaluation yields the matrices (2.5)–(2.8). (Note that Σ_m is rational also for $m = 5$, although $\operatorname{Re}\omega_5 = (\sqrt{5}-1)/4$ is not. This is easily explained by Galois theory.)

The case $m = 2$, with $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, is a special case of the so-called Friedman's urn [20] (studied already by Bernstein [4]), and the result follows alternatively directly from Bernstein [4], [5] or Freedman [19]; see also [25, Example 3.27].

In Case (ii), $m = 6$, we similarly find by [25, Lemma 5.3(iv)]

$$\Sigma_6 = v_1 v_1^* + v_5 v_5^* = 6^{-2}\bigl(2\operatorname{Re}\omega_6^{j-k}\bigr)_{j,k},$$

which yields (2.9).

In Case (iii), (2.3) follows from [25, Theorem 3.24], with $d = 0$ and $\lambda_2 = \omega_m = \alpha+\mathrm{i}\beta$. Let us now denote the eigenvectors u_j and v_j belonging to ω^j by u_{ω^j} and v_{ω^j}. Then [25, Theorem 3.24] further shows that $\widetilde{W}_m$ belongs to the linear span E_ω of $v_\omega = (m^{-1}\omega^{-j})_j$; hence $\widetilde{W}_m = \bigl(\widetilde{Z}_m\omega^{-j}\bigr)_{j=0}^{m-1}$ as asserted for some complex random $\widetilde{Z}_m := m^{-1}u_\omega \cdot \widetilde{W}_m$. By [25, Theorem 3.26], (2.10) holds with $Z = u_\omega \cdot W_\omega$, and W_ω as in [25, Theorem 3.1].

Recall that we start the urn with a single ball with label 0. Let, as in [25, Theorem 3.9], $W_{\omega,i}$ be the limit random variable corresponding to W_ω if we instead start with a single ball of type i. By symmetry, $W_{\omega,i}$ is obtained from $W_{\omega,0} = W_\omega$ by a cyclic shift of the components, and thus $Z_i := u_\omega \cdot W_{\omega,i} = \omega^i Z$. By [25, Theorem 3.9(ii)],

$$W_{\omega,i} \overset{\mathrm{d}}{=} U^\omega(W_{\omega,i} + \omega W'_{\omega,i+1}) \tag{3.2}$$

with $W'_{\omega,i+1}$ distributed as $W_{\omega,i+1}$ and independent of $W'_{\omega,i}$. Consequently,

$$Z_i \overset{\mathrm{d}}{=} U^\omega(Z_i + Z'_{i+1}) \overset{\mathrm{d}}{=} U^\omega(Z_i + \omega Z'_i),$$

with Z_i, Z'_{i+1}, Z'_i independent, and taking $i = 0$ we obtain (2.11). Moreover, $\mathbb{E}\, Z = u_{\omega 0} = 1$ by [25, Theorem 3.10].

Conversely, (2.11) implies (3.2) with $W_{\omega,i} = \omega^i Z v_\omega$, and thus [25, Theorem 3.9(iii)] implies that the distribution of Z is determined by (2.11) and $\mathbb{E}\, Z$. The distribution of $\widetilde{Z}$ then is determined by (2.10).

For higher moments of Z, we take moments in (2.11). For example,

$$\mathbb{E}\, Z^2 = \mathbb{E}\, U^{2\omega}\, \mathbb{E}(Z + \omega Z')^2 = (1+2\omega)^{-1}\bigl((1+\omega^2)\,\mathbb{E}\, Z^2 + 2\omega(\mathbb{E}\, Z)^2\bigr),$$

which, using $\mathbb{E}\, Z = 1$, gives $\mathbb{E}\, Z^2 = 2/(2-\omega)$ after rearrangement. We leave the corresponding calculations for $\mathbb{E}\,|Z|^2$ and $\mathbb{E}\, Z^3$ to the reader. See also [25, Theorem 3.10].

The formulas for moments of $\widetilde{Z}$ then follow from (2.10), using $\mathbb{E}\, W^z = \Gamma(1+z)$ when $\operatorname{Re} z > -1$; see also [25, Theorem 3.26] and [23, Section 2].

4. Binary search trees

To describe the profile of the random BST in terms of an urn model, we make a simple transformation. A BST with n vertices has $n+1$ possible positions for a new vertex. We augment the tree by adding $n+1$ new vertices at these positions; the $n+1$ new vertices are called *external* and the n original vertices are called *internal*. Thus every internal vertex has two children, and every external vertex has none.

Thus, $(X_j)_{j=0}^{\infty}$ is now the profile of the internal vertices. We similarly define $(Y_j)_{j=0}^{\infty}$ as the profile of the external vertices, and note that, since every internal vertex has exactly two children,

$$2X_{j-1} = X_j + Y_j, \qquad j \geq 1.$$

For $j = 0$ we instead have $X_0 + Y_0 = X_0 = 1$. (Recall that the root has no parent.) Passing to congruence classes modulo m we thus have, for every j,

$$2X_{j-1}^{(m)} - X_j^{(m)} = Y_j^{(m)} - \delta_{j0}. \tag{4.1}$$

The growth of the augmented tree can be described as follows: Choose an external vertex at random, convert it to an internal vertex and add two new external vertices as its children. The distribution of depths for external vertices modulo m is thus given by a generalized Pólya urn similar to the one in Section 3, with the difference that when we draw a ball with label j, we remove it and add two balls with label $j+1$. (We start with 2 balls with label 1; alternatively, we start with a single ball with label 0 and make one more draw.)

The matrix A is now $(2\delta_{i,j+1} - \delta_{i,j})_{i,j=0}^{m-1}$, again with index addition modulo m. The eigenvectors are the same u_j and v_j as in Section 3, but the corresponding eigenvalue is now $2\omega^j - 1$. In particular, the largest eigenvalue (i.e., the one with largest real part) is $\lambda_1 = 1$ (as before), and the second largest are $\lambda_2 = 2\omega - 1$ and $\overline{\lambda_2}$, with real part $\operatorname{Re}\lambda_2 = 2\cos(2\pi/m) - 1$.

Hence, the condition $\operatorname{Re}\lambda_2 < 1/2$ becomes $\cos(2\pi/m) < 3/4$, which holds for $m \leq 8$, while for $m \geq 9$ we have $\cos(2\pi/m) > 3/4$ and thus $\operatorname{Re}\lambda_2 > 1/2$.

We now obtain, exactly as in Section 3, normal convergence of $\mathbf{Y}^{(m)} = (Y_0^{(m)}, \dots, Y_{m-1}^{(m)})$ when $m \leq 8$. More precisely, by [25, Theorem 3.22 and Lemma 5.3(iii) or Lemmas 5.3(ii) and 5.4], (2.2) holds for $\mathbf{Y}^{(m)}$ with

$$\Sigma_m = \sum_{j=1}^{m-1} \frac{|2\omega^j - 1|^2}{3 - 4\operatorname{Re}\omega^j} v_j v_j^* = \sum_{j=1}^{m-1} \frac{5 - 4\operatorname{Re}\omega^j}{3 - 4\operatorname{Re}\omega^j} v_j v_j^*. \tag{4.2}$$

Similarly, by [25, Theorem 3.24], when $m \geq 9$, (2.3) and (2.4) hold for $\mathbf{Y}^{(m)}$, for some $\widetilde{Z}_m$ and $\widetilde{W}_m = \bigl(\widetilde{Z}_m \omega^{-j}\bigr)_{j=0}^{m-1}$. Further, by [25, Theorems 3.26, 3.9 and 3.10],

$$Z \overset{\mathrm{d}}{=} \tfrac{m}{2} W^{2\omega-1} \widetilde{Z}_m, \tag{4.3}$$

where Z satisfies (2.17) and $\mathbb{E}\, Z = 1$.

To obtain the results for $\mathbf{X}^{(m)}$, we use (4.1). It is convenient to solve this convolution equation by taking the Fourier transform, which yields, for all k,

$$(2\omega^k - 1)\widehat{X}^{(m)}(k) = \widehat{Y}^{(m)}(k) - 1. \tag{4.4}$$

In the case $m \leq 8$, we have in analogy with Remark 2.6, $\widehat{Y}^{(m)}(0) = n + 1$ and the joint convergence

$$n^{-1/2}\widehat{Y}^{(m)}(k) \overset{\mathrm{d}}{\longrightarrow} V_k, \qquad k = 1, \dots, m-1, \tag{4.5}$$

where V_k are complex Gaussian variables satisfying (i)–(iii) and (v)–(vii) in Remark 2.6 together with

$$\text{(iv}''''\text{)}\ \mathbb{E}\,|V_k|^2 = \frac{|2\omega_m^k - 1|^2}{3 - 4\operatorname{Re}\omega_m^k} = \frac{5 - 4\operatorname{Re}\omega_m^k}{3 - 4\operatorname{Re}\omega_m^k} = \frac{5 - 4\cos(2\pi k/m)}{3 - 4\cos(2\pi k/m)}.$$

It follows immediately from (4.4) and (4.5) that

$$n^{-1/2}\widehat{X}^{(m)}(k) \xrightarrow{\text{d}} (2\omega^k - 1)^{-1}V_k, \qquad k = 1, \dots, m-1, \tag{4.6}$$

which yields the statement in Remark 2.11 (for $m \le 8$), with the meaning of V_k changed.

Similarly, for $m \ge 9$, we have, with $\lambda := 2\omega - 1$, in analogy with (2.13), $n^{-\lambda}\widehat{Y}^{(m)}(1) \xrightarrow{\text{a.s.}} \frac{m}{2}\widetilde{Z}_m$, and thus

$$n^{-\lambda}\widehat{X}^{(m)}(1) = n^{-\lambda}\lambda^{-1}\widehat{Y}^{(m)}(1) + o(1) \xrightarrow{\text{a.s.}} \lambda^{-1}\frac{m}{2}\widetilde{Z}_m,$$

while $n^{-\lambda}\widehat{X}^{(m)}(k) \xrightarrow{\text{a.s.}} 0$ for $k = 2, \dots, m-2$. We change the meaning of $\widetilde{Z}_m$ (replacing $\widetilde{Z}_m$ by $\lambda\widetilde{Z}_m$) and write this as (2.18), simultaneously changing (4.3) to (2.16).

Theorem 2.7 now follows by taking the inverse Fourier transform. When $m \le 8$, we obtain (2.2) with

$$\Sigma_m = \sum_{j=1}^{m-1} \frac{1}{3 - 4\operatorname{Re}\omega_m^j} v_j v_j^*, \tag{4.7}$$

which gives the explicit values in Remark 2.8 (with some help of `Maple`).

Remark 4.1. The covariance matrices in (4.2) for the case of external vertices are circulant and explicitly given by the following first rows:

$$\begin{aligned}
m = 2: &\quad \tfrac{1}{28}(9, -9),\\
m = 3: &\quad \tfrac{1}{45}(14, -7, -7),\\
m = 4: &\quad \tfrac{1}{336}(97, -27, -43, -27),\\
m = 5: &\quad \tfrac{1}{275}(76, -9, -29, -29, -9),\\
m = 6: &\quad \tfrac{1}{1260}(353, 11, -109, -157, -109, 11),\\
m = 7: &\quad \tfrac{1}{637}(202, 41, -43, -99, -99, -43, 41),\\
m = 8: &\quad \tfrac{1}{1344}(685, 309, -43, -363, -491, -363, -43, 309).
\end{aligned}$$

5. Conditioned Galton–Watson trees

For Galton–Watson trees, we use generating functions and singularity analysis. See [6] for similar arguments. Given a tree T, we define its profile polynomial by

$$S(x) = S(x;T) := \sum_{v \in T} x^{d(v)} = \sum_j x^j X_j(T). \tag{5.1}$$

We will first find the asymptotic distribution of $S(x;T_n)$ for x on the unit circle (excluding the trivial case $x=1$). Note that $S(x)$ for $|x|=1$ is the Fourier transform of the sequence (X_j) as a function on $\mathbb{Z}$.

Letting $\mathcal{T}$ be a random Galton–Watson tree, we define the generating functions, for $k\geq 0$,

$$F_k(t;x_1,\dots,x_k):=\mathbb{E}\Bigl(t^{|\mathcal{T}|}\prod_{i=1}^k S(x_i;\mathcal{T})\Bigr). \tag{5.2}$$

Here t and $x_1,\dots,x_k$ are complex numbers. (It is also possible to regard F_k as a formal power series, but we will need analytic functions.) We regard $x_1,\dots,x_k$ as fixed and consider F_k as a function of t. We consider only x_i with $|x_i|\leq 1$; then $|S(x_i;\mathcal{T})|\leq|\mathcal{T}|$ and the expectation in (5.2) exists at least for $|t|<1$. Thus (5.2) defines F_k as an analytic function of t in the unit disc $|t|<1$. We will soon see that it can be continued to a larger domain.

Let D_0 be the degree of the root. If we condition on $D_0=q\geq 0$, then the random tree $\mathcal{T}$ consists of the root plus q branches $T_1,\dots,T_q$ that are independent and have the same distribution as $\mathcal{T}$. Further, $|\mathcal{T}|=1+\sum_1^q|T_j|$ and $S(x;\mathcal{T})=1+\sum_{j=1}^q xS(x;T_j)$, and thus, summing over all sequences $I_0,\dots,I_q$ of disjoint (possibly empty) subsets of $\{1,\dots,k\}$ with $\bigcup_1^q I_j=\{1,\dots,k\}$,

$$\prod_{i=1}^k S(x_i;\mathcal{T})=\sum_{I_0,\dots,I_q}\prod_{j=1}^q\prod_{i\in I_j}x_iS(x_i;T_j).$$

Consequently,

$$\begin{aligned}\mathbb{E}\Bigl(t^{|\mathcal{T}|}\prod_{i=1}^k S(x_i;\mathcal{T})\Bigm| D_0=q\Bigr)&=\mathbb{E}\sum_{I_0,\dots,I_q}t\prod_{j=1}^q\Bigl(t^{|T_j|}\prod_{i\in I_j}x_iS(x_i;T_j)\Bigr)\\&=t\sum_{I_0,\dots,I_q}\Bigl(\prod_{i\notin I_0}x_i\Bigr)\prod_{j=1}^q F_{|I_j|}(t;\{x_i,\,i\in I_j\}).\end{aligned}$$

The terms in the latter sum do not depend on the order of $I_1,\dots,I_q$. Thus, if $\sum^*_{I_0,\dots,I_l}$ denotes the sum over such sequences $I_0,\dots,I_l$ with $I_1,\dots,I_l$ non-empty and in, say, lexicographic order, then, with $q^{\underline{l}}:=q(q-1)\cdots(q-l+1)$,

$$\begin{aligned}&\mathbb{E}\Bigl(t^{|\mathcal{T}|}\prod_{i=1}^k S(x_i;\mathcal{T})\Bigm| D_0=q\Bigr)\\&\qquad=t\sum_{l=0}^k\sum^*_{I_0,\dots,I_l}q^{\underline{l}}\Bigl(\prod_{i\notin I_0}x_i\Bigr)\prod_{j=1}^l F_{|I_j|}(t;\{x_i,\,i\in I_j\})F_0(t)^{q-l}. \end{aligned}\tag{5.3}$$

Now take the expectation, i.e. multiply by $\mathbb{P}(D_0 = q)$ and sum over q. We have, for $|z| < 1$ at least, since $D_0 \overset{\mathrm{d}}{=} \xi$ and thus $\mathbb{E}\, z^{D_0} = \varphi(z)$,

$$\sum_{q \geq l} q^{\underline{l}} z^{q-l}\, \mathbb{P}(D_0 = q) = \mathbb{E}\big(D_0^{\underline{l}} z^{D_0 - l}\big) = \varphi^{(l)}(z),$$

and thus (5.3) yields

$$F_k(t; x_1, \dots, x_k) = t \sum_{l=0}^{k} \sum_{I_0, \dots, I_l}^{*} \varphi^{(l)}(F_0(t)) \Big(\prod_{i \notin I_0} x_i\Big) \prod_{j=1}^{l} F_{|I_j|}(t; \{x_i,\, i \in I_j\}).$$

In particular, $k = 0$ yields the well-known formula $F_0(t) = t\varphi(F_0(t))$. The next two cases are

$$\begin{aligned} F_1(t; x) &= t\varphi(F_0(t)) + t\varphi'(F_0(t)) x F_1(t; x), \\ F_2(t; x, y) &= t\varphi(F_0(t)) + t\varphi'(F_0(t))\big(x F_1(t; x) + y F_1(t; y) + xy F_2(t; x, y)\big) \\ &\qquad + t\varphi''(F_0(t)) xy F_1(t; x) F_1(t; y). \end{aligned}$$

We thus have, recalling $t\varphi(F_0(t)) = F_0(t)$,

$$F_1(t; x) = \frac{F_0(t)}{1 - xt\varphi'(F_0(t))}, \tag{5.4}$$

$$\begin{aligned} F_2(t; x, y) = \big(1 - xyt\varphi'(F_0(t))\big)^{-1} \Big(& F_0(t) + t\varphi'(F_0(t))\big(x F_1(t; x) + y F_1(t; y)\big) \\ & + t\varphi''(F_0(t)) xy F_1(t; x) F_1(t; y)\Big), \end{aligned} \tag{5.5}$$

and in general

$$\begin{aligned} F_k(t; x_1, \dots, x_k) &= \Big(1 - t\Big(\prod_{i=1}^{k} x_i\Big)\varphi'(F_0(t))\Big)^{-1} \\ &\quad \times t \sum_{l=0}^{k} \sum_{I_0, \dots, I_l}^{**} \varphi^{(l)}(F_0(t)) \Big(\prod_{i \notin I_0} x_i\Big) \prod_{j=1}^{l} F_{|I_j|}(t; \{x_i,\, i \in I_j\}), \end{aligned} \tag{5.6}$$

where $\sum^{**}$ means $\sum^{*}$ with the single term with $l = 1$ and $|I_1| = k$ omitted. This gives recursively an explicit formula for each $F_k(t; x_1, \dots, x_k)$ as a rational function of $x_1, \dots, x_k$ and $t\varphi^{(l)}(F_0(t))$, $0 \leq l \leq k$.

We say, see [18, Chapter VI], that a Δ-*domain* is a domain of the type $\{z : |z| < 1 + \varepsilon,\ |\arg(z - 1)| > \pi/2 - \delta\}$ for some (small) positive ε and δ, and that a function is Δ-*analytic* if it is analytic in some Δ-domain, or can be extended to such a function.

Suppose for simplicity in the sequel that ξ is aperiodic. (The periodic case is similar with standard modifications as in [18, Chapter VI.7]; we omit the details.) Then, by a standard result in singularity analysis, see e.g. [18, Proposition VI.1], $F_0(t)$ is Δ-analytic, with

$$F_0(t) = 1 - \sqrt{2/\sigma^2}(1 - t)^{1/2} + O(1 - t) \qquad \text{as } t \to 1. \tag{5.7}$$

(Here and below, we consider only t in a suitable Δ-domain.) It follows that in a (possibly smaller) Δ-domain, $|F_0(t)| < 1$ and hence $|\varphi'(F_0(t))| < \varphi'(1) = 1$. Further,

$$\begin{aligned}\varphi'(F_0(t)) &= 1 + \varphi''(1)\bigl(F_0(t) - 1\bigr) + O\bigl(F_0(t) - 1\bigr)^2 \\ &= 1 - \sqrt{2}\sigma(1-t)^{1/2} + O(1-t),\end{aligned} \tag{5.8}$$

and it follows easily that $|t\varphi'(F_0(t))| < 1$ in a Δ-domain. Hence, for every fixed x with $|x| \le 1$, $\bigl(1 - xt\varphi'(F_0(t))\bigr)^{-1}$ is Δ-analytic, with

$$\bigl(1 - xt\varphi'(F_0(t))\bigr)^{-1} = \begin{cases} 2^{-1/2}\sigma^{-1}(1-t)^{-1/2} + O(1), & \text{if } x = 1, \\ O(1), & \text{if } x \ne 1. \end{cases} \tag{5.9}$$

Lemma 5.1. *For $k \ge 1$ and any complex $x_1, \dots, x_k$ with $|x_i| \le 1$ and $x_i \ne 1$, $F_k(t; x_1, \dots, x_k)$ is Δ-analytic (as a function of t), with*

$$F_k(t; x_1, \dots, x_k) = a_k(x_1, \dots, x_k)(1-t)^{-(k-1)/2} + O\bigl((1-t)^{-(k-2)/2}\bigr), \tag{5.10}$$

where $a_1(x) = 1/(1-x)$, and for $k \ge 2$, $a_k(x_1, \dots, x_k) = 0$ if $x_1 \cdots x_k \ne 1$, while if $x_1 \cdots x_k = 1$,

$$a_k(x_1, \dots, x_k) = 2^{-3/2}\sigma \sum_{\emptyset \subsetneq I \subsetneq \{1,\dots,k\}} a_{|I|}(x_i : i \in I) a_{k-|I|}(x_i : i \notin I). \tag{5.11}$$

In particular, when $|x| = 1$, $a_2(x, \bar{x}) = 2^{-1/2}\sigma|1-x|^{-2}$.

Proof. The Δ-analyticity follows by (5.6) and induction, using the results just shown.

For $k = 1$, the expansion (5.10) follows from (5.4), (5.7) and (5.8).

Similarly, (5.10) for $k = 2$ follows from (5.5) together with (5.7), (5.8) and (5.9); this also yields $a_2(x, y) = 0$ if $xy \ne 1$ and for $xy = 1$, using also $\varphi''(1) = \sigma^2$,

$$a_2(x, y) = 2^{-1/2}\sigma^{-1}\bigl(F_0(1) + xa_1(x) + ya_1(y) + \sigma^2 xya_1(x)a_1(y)\bigr),$$

which equals $2^{-1/2}\sigma|1-x|^{-2}$ because now $|x| = |y| = 1$ and $y = \bar{x}$ and thus

$$xa_1(x) + ya_1(y) = \frac{x}{1-x} + \frac{y}{1-y} = 2\operatorname{Re}\Bigl(\frac{x}{1-x}\Bigr) = -1.$$

For $k \ge 3$ we argue similarly. By induction, all terms in the sum in (5.6) are $O\bigl((1-t)^{-(k-2)/2}\bigr)$. The result when $x_1 \cdots x_k \ne 1$ follows immediately by (5.9).

Assume now $x_1 \cdots x_k = 1$. If $|I_1| = k-1$, then $I_1 = \{1, \dots, k\} \setminus x_p$ for some p, and thus $\prod_{i \in I_1} x_i = x_p^{-1} \ne 1$; hence by induction $F_{k-1}(t; \{x_i : i \in I_1\}) = O\bigl((1-t)^{-(k-3)/2}\bigr)$. The leading terms in (5.6) are thus those with $l = 2$ and $I_0 = \emptyset$, $I_2 = \{1, \dots, k\} \setminus I_1$, which proves the claim, including (5.11), by another application of (5.9). (Note the factor 1/2 because we assume that I_1 and I_2 are in order, but not necessarily I and its complement.) $\square$

We next solve the recursion (5.11). Let $N(x_1,\dots,x_k)$ be the number of pairings of $x_1,\dots,x_k$ into $k/2$ pairs of the type $\{x,\bar{x}\}$. (Thus $N(x_1,\dots,x_k)=0$ if k is odd.)

Lemma 5.2. *Let $k\ge 2$. Suppose that $x_1,\dots,x_k\in\{x\in\mathbb{C}:|x|=1$ but $x\neq 1\}$. Then*

$$a_k(x_1,\dots,x_k)=\frac{\Gamma\bigl((k-1)/2\bigr)}{\sqrt{2\pi}\sigma}N(x_1,\dots,x_k)\prod_{i=1}^{k}\frac{\sigma}{1-x_i}.$$

Proof. For $k=2$, the result follows directly from Lemma 5.1.

For $k\ge 3$, we use induction. The result is trivial if $x_1\cdots x_k\neq 1$. Hence, we assume $x_1\cdots x_k=1$ and use (5.11). First, note that it suffices to consider I with $|I|$ even. In fact, if $|I|$ is odd and $|I|\ge 3$, then $N_{|I|}(x_i:i\in I)=0$, and thus $a_{|I|}(x_i:i\in I)=0$ by induction. If $|I|=1$, then $I=\{x_p\}$ for some p. Since $\prod_{i\notin I_1}x_i=x_p^{-1}\neq 1$, we have $N_{k-|I|}(x_i:i\notin I)=0$, and thus by induction $a_{k-|I|}(x_i:i\notin I)=0$. Similarly, we may assume that $k-|I|$ is even.

The result thus holds when k is odd (with $a_k(x_1,\dots,x_k)=0$).

Now let k be even, $k=2l$ with $l\ge 2$. We use induction on the right hand side of (5.11). Say that a pairing of $x_1,\dots,x_k$ is *good* if each pair consist of two conjugate numbers. Note that $N_{|I|}(x_i:i\in I)N_{k-|I|}(x_i:i\notin I)$ equals the number of pairs (α,β) where α is a good pairing of $(x_i:i\in I)$ and β is a good pairing of $(x_i:i\notin I)$. Each such pair (α,β) defines a good pairing of $x_1,\dots,x_k$, and conversely, each good pairing of $x_1,\dots,x_k$ splits into good pairings α and β in 2^l-2 ways: $\binom{l}{j}$ ways with $|I|=2j$ for each $j=1,\dots,l-1$. Consequently, (5.11) yields,

$$a_k(x_1,\dots,x_k)=2^{-3/2}\sigma\sum_{j=1}^{l-1}\binom{l}{j}N(x_1,\dots,x_k)\times\frac{\Gamma(j-1/2)\Gamma(l-j-1/2)}{2\pi\sigma^2}\prod_{i=1}^{k}\frac{\sigma}{1-x_i}.$$

To complete the induction step, it is now sufficient to verify

$$\sum_{j=1}^{l-1}\binom{l}{j}\Gamma(j-1/2)\Gamma(l-j-1/2)=4\sqrt{\pi}\Gamma(l-1/2),$$

for $l\ge 2$. This is an immediate consequence of the binomial convolution

$$\sum_{j=0}^{l}\binom{l}{j}\Gamma(j-1/2)\Gamma(l-j-1/2)=l!\sum_{j=0}^{l}\frac{\Gamma(j-1/2)}{j!}\frac{\Gamma(l-j-1/2)}{(l-j)!}$$
$$=l!\sum_{j=0}^{l}(-1)^l\Gamma(-1/2)^2\binom{1/2}{j}\binom{1/2}{l-j}=l!\,(-1)^l\Gamma(-1/2)^2\binom{1}{l}=0,$$

when $l \geq 2$, since the terms with $j = 0$ and $j = l$ both are $\Gamma(-1/2)\Gamma(l-1/2) = -2\sqrt{\pi}\Gamma(l-1/2)$. □

Lemma 5.3. *Let $k \geq 1$. Suppose that $x_1, \dots, x_k \in \{x \in \mathbb{C} : |x| = 1$ but $x \neq 1\}$. Then*

$$\mathbb{E}\big(S(x_1;T_n)\cdots S(x_k;T_n)\big) = N(x_1,\dots,x_k)\Big(\prod_{i=1}^{k}\frac{\sigma}{1-x_i}\Big)n^{k/2} + O\big(n^{(k-1)/2}\big).$$

Proof. By (5.4),

$$\mathbb{E}\big(S(x_1;T_n)\cdots S(x_k;T_n)\big) = \frac{[t^n]F_k(t;x_1,\dots,x_k)}{\mathbb{P}(|\mathcal{T}|=n)} = \frac{[t^n]F_k(t;x_1,\dots,x_k)}{[t^n]F_0(t)}.$$

We obtain from (5.10) by standard singularity analysis, see e.g. [18, Chapter VI], that for $k \geq 2$,

$$[t^n]F_k(t;x_1,\dots,x_k) = \frac{a_k(x_1,\dots,x_k)}{\Gamma\big((k-1)/2\big)}n^{(k-3)/2} + O\big(n^{(k-4)/2}\big),$$

while for $k = 1$,

$$[t^n]F_1(t;x) = O\big(n^{-3/2}\big).$$

Similarly, (5.7) yields, as is well-known,

$$[t^n]F_0(t) = \frac{1}{\sqrt{2\pi}\sigma}n^{-3/2} + O\big(n^{-5/2}\big).$$

The result follows from these formulas and Lemma 5.2. □

We can now identify the asymptotic moments and thus the asymptotic distribution of $S(x;T_n)$.

Theorem 5.4. *Let $U(x)$ be a family of complex Gaussian random variables, defined for $|x| = 1$ but $x \neq 1$, such that*

(i) *$U(x)$ is symmetric complex Gaussian when $\operatorname{Im} x \neq 0$, with $\mathbb{E}\,|U(x)|^2 = 1/|1-x|^2$;*
(ii) *$U(x)$ is symmetric real Gaussian when $\operatorname{Im} x = 0$ (i.e., when $x = -1$), with $\mathbb{E}\,|U(x)|^2 = 1/|1-x|^2$;*
(iii) *$U(\bar{x}) = \overline{U(x)}$;*
(iv) *the variables $U(x)$, $\operatorname{Im} x \geq 0$, are independent.*

Then, for the CGWT, $S(x;T_n)/\sqrt{n} \xrightarrow{\text{d}} \sigma U(x)$, jointly for all such x.

Proof. Note that the assumptions imply that $\mathbb{E}\,U(x)U(y) = 1/|1-x|^2$ if $xy = 1$ (and thus $y = \bar{x}$), but $\mathbb{E}\,U(x)U(y) = 0$ otherwise. By Lemma 5.3 and the formula [24, Theorem 1.28] for joint moments of Gaussian variables (known as Wick's theorem),

$$n^{-k/2}\,\mathbb{E}\big(S(x_1;T_n)\cdots S(x_k;T_n)\big) \to \sigma^k\,\mathbb{E}\big(U(x_1)\cdots U(x_k)\big).$$

Replacing one or several x_i by their conjugates, we see that the same holds if we replace some S and U by their conjugates. Hence the result holds by the method of moments (applied to the real and imaginary parts). □

Proof of Theorem 2.12. As remarked above, for any tree T, the numbers $S(x;T)$ for $|x| = 1$ form the Fourier transform of the sequence $(X_j(T))$. It follows that the discrete Fourier transform $\widehat{\mathbf{X}}^{(m)}$ of $\mathbf{X}^{(m)}(T)$ equals the vector $(S(\omega^k;T))_{k=0}^{m-1}$, where $\omega = \omega_m$, cf. (2.1) and (5.1). Hence, by Fourier inversion,

$$X_j(T) = \frac{1}{m}\sum_{k=0}^{m-1} \omega^{-jk} S(\omega^k;T).$$

Theorem 5.4 thus implies, since trivially $S(1;T_n) = n$,

$$n^{-1/2}\Big(X_j(T_n) - \frac{n}{m}\Big) \xrightarrow{\mathrm{d}} Z_j := \frac{\sigma}{m}\sum_{k=1}^{m-1} \omega^{-jk} U(\omega^k),$$

jointly for all j. Since the variables $U(\omega^k)$ are jointly (complex) Gaussian, the variables Z_j are too; moreover, each Z_j is real. Clearly, $\mathbb{E}\, Z_j = 0$, and the covariance matrix is given by

$$\begin{aligned}\mathbb{E}(Z_i Z_j) = \mathbb{E}(Z_i \overline{Z_j}) &= \frac{\sigma^2}{m^2}\sum_{k=1}^{m-1} \omega^{(j-i)k}\, \mathbb{E}\,|U(\omega^k)|^2 \\ &= \frac{\sigma^2}{m^2}\sum_{k=1}^{m-1} \omega^{(j-i)k} |1-\omega^k|^{-2}. \end{aligned} \tag{5.12}$$

To evaluate this sum, define a function f on the group $\mathbb{Z}_m = \{0,\dots,m-1\}$ by $f(j) = j - (m-1)/2$. Then its Fourier transform is

$$\widehat{f}(k) = \sum_{j=0}^{m-1}\big(j - \frac{m-1}{2}\big)\omega^{jk} = \frac{m}{\omega^k - 1}, \qquad k \neq 0,$$

while $\widehat{f}(0) = 0$.

Let further $g := f * \check{f}$ on $\mathbb{Z}_m$, where $\check{f}(i) = f(-i) = f(m-i)$, i.e. $g(j) = \sum_{i\in\mathbb{Z}_m} f(j+i)f(i)$, with $j+i$ taken modulo m. Then $\widehat{g}(k) = |\widehat{f}(k)|^2 = m^2|\omega^k - 1|^{-2}$, $k \neq 0$, and thus, by Fourier inversion again,

$$\sum_{k=1}^{m-1} \omega^{(j-i)k}|1-\omega^k|^{-2} = m^{-2}\sum_{k=0}^{m-1} \omega^{(j-i)k}\widehat{g}(k) = m^{-1} g(i-j).$$

Hence, by (5.12), $\mathbb{E}(Z_i Z_j) = \sigma^2 m^{-3} g(i-j)$. It remains only to evaluate g. For $0 \le j < m$,

$$\begin{aligned} g(j) &= \sum_{i=0}^{m-1-j}(j+i)i + \sum_{i=m-j}^{m-1}(j+i-m)i - m\Big(\frac{m-1}{2}\Big)^2 \\ &= m\frac{m^2-1}{12} - m\frac{j(m-j)}{2}. \end{aligned}$$

□

6. Oscillations

In Case (iii) of Theorem 2.1 ($m \geq 7$), we do not have convergence in distribution: the sequence of random vectors $n^{-\alpha}\big(\mathbf{X}^{(m)} - \frac{n}{m}\mathbf{1}\big)$, $n = 1, 2, \dots$, is tight and thus suitable subsequences converge as is shown explicitly in (2.4), but different subsequences may have different limits, and thus $n^{-\alpha}\big(X_0^{(m)} - \frac{n}{m}\mathbf{1}\big)$, for example, does not have a limit distribution. Indeed, suppose that $n^{-\alpha}\big(X_0^{(m)} - \frac{n}{m}\mathbf{1}\big) \xrightarrow{\mathrm{d}} V$, say. Then, by (2.4), $\operatorname{Re}(e^{\mathrm{i}\gamma}\widetilde{Z}_m) \overset{\mathrm{d}}{=} V$ for every $\gamma \in [0, 2\pi]$. In particular, $\operatorname{Re}(e^{\mathrm{i}\gamma}\,\mathbb{E}\,\widetilde{Z}_m) = \mathbb{E}\operatorname{Re}(e^{\mathrm{i}\gamma}\widetilde{Z}_m) = \mathbb{E}\,V$ is independent of $\gamma \in [0, 2\pi]$, which is a contradicion because $\mathbb{E}\,\widetilde{Z}_m \neq 0$ by Remark 2.3.

Nevertheless, it is conceivable (although implausible) that $X_0^{(m)}$ has a limit distribution if we choose the norming constants carefully, i.e. that $a_n(X_0^{(m)} - b_n) \xrightarrow{\mathrm{d}} V$ for some non-degenerate V and real constants $a_n > 0$ and b_n. It then would follow from (2.4) that $\operatorname{Re}(e^{\mathrm{i}\gamma}\widetilde{Z}_m) \overset{\mathrm{d}}{=} a_\gamma V + b_\gamma$ for every $\gamma \in [0, 2\pi]$ and some real constants $a_\gamma \geq 0$ and b_γ, see e.g. [22, Section 9.2]. In other words, $\operatorname{Re}(e^{\mathrm{i}\gamma}\widetilde{Z}_m)$ would have a distribution of the same type for every γ, except when it is degenerate. We conjecture that this does not happen.

Conjecture 6.1. *For RRT and every $m \geq 7$, there are oscillations in Theorem 2.1(iii); $a_n(X_0^{(m)} - b_n)$ does not have a non-degenerate limit distribution for any sequence of norming constants $a_n \geq 0$ and b_n. The same holds for BST and every $m \geq 9$.*

We can verify this conjecture in many cases by a general result.

Proposition 6.2. *Let Z be a complex random variable such that $\mathbb{E}\,|Z|^3 < \infty$. Suppose that there exists a random variable V and, for every $\gamma \in [0, 2\pi]$, some real constants $a_\gamma \geq 0$ and b_γ such that $\operatorname{Re}\big(e^{\mathrm{i}\gamma}Z\big) \overset{\mathrm{d}}{=} a_\gamma V + b_\gamma$. Then either*

(i) *$Z \overset{\mathrm{d}}{=} aW + b$ for some real random variable W and some complex constants a and b, and thus $|\mathbb{E}(Z - \mathbb{E}\,Z)^2| = \mathbb{E}\,|Z - \mathbb{E}\,Z|^2$; or*

(ii) $\mathbb{E}(Z - \mathbb{E}\,Z)^3 = \mathbb{E}(Z - \mathbb{E}\,Z)^2\overline{(Z - \mathbb{E}\,Z)} = 0$.

Proof. If all $a_\gamma = 0$, then $\operatorname{Re}\big(e^{\mathrm{i}\gamma}(Z - \mathbb{E}\,Z)\big) = 0$ a.s. for every γ, and it follows by the Cramér–Wold device that $Z = \mathbb{E}\,Z$ a.s., a special case of (i).

Thus assume that some $a_\gamma > 0$. Then $\mathbb{E}\,|V|^3 < \infty$. By replacing Z by $Z - \mathbb{E}\,Z$ and V by $V - \mathbb{E}\,V$ (changing b_γ accordingly), we may assume that $\mathbb{E}\,Z = \mathbb{E}\,V = 0$, and thus $b_\gamma = 0$. If $V = 0$ a.s., then, by the Cramér–Wold device again, $Z = 0$ a.s., and (i) holds. Assume thus $\mathbb{E}\,V^2 > 0$; rescaling V we may assume $\mathbb{E}\,V^2 = 1$. Define

$$X_\gamma := \operatorname{Re}\big(e^{\mathrm{i}\gamma}Z\big) \overset{\mathrm{d}}{=} a_\gamma V,$$
$$Z_\gamma := 2e^{\mathrm{i}\gamma}X_\gamma = e^{2\mathrm{i}\gamma}Z + \bar{Z}.$$

We have

$$\mathbb{E}\, Z_\gamma^2 = 4e^{2\mathrm{i}\gamma}\,\mathbb{E}\, X_\gamma^2 = 4e^{2\mathrm{i}\gamma} a_\gamma^2,$$
$$\mathbb{E}\, Z_\gamma^3 = 8e^{3\mathrm{i}\gamma}\,\mathbb{E}\, X_\gamma^3 = 8e^{3\mathrm{i}\gamma} a_\gamma^3\,\mathbb{E}\, V^3,$$

and thus

$$\bigl(\mathbb{E}\, Z_\gamma^3\bigr)^2 = (\mathbb{E}\, V^3)^2\bigl(\mathbb{E}\, Z_\gamma^2\bigr)^3. \tag{6.1}$$

On the other hand,

$$\mathbb{E}\, Z_\gamma^2 = \mathbb{E}\bigl(e^{2\mathrm{i}\gamma} Z + \bar{Z}\bigr)^2 = e^{4\mathrm{i}\gamma}\,\mathbb{E}\, Z^2 + 2e^{2\mathrm{i}\gamma}\,\mathbb{E}\,|Z|^2 + \mathbb{E}\,\bar{Z}^2, \tag{6.2}$$
$$\mathbb{E}\, Z_\gamma^3 = e^{6\mathrm{i}\gamma}\,\mathbb{E}\, Z^3 + 3e^{4\mathrm{i}\gamma}\,\mathbb{E}(Z^2\bar{Z}) + 3e^{2\mathrm{i}\gamma}\,\mathbb{E}(Z\bar{Z}^2) + \mathbb{E}\,\bar{Z}^3. \tag{6.3}$$

Hence, $\mathbb{E}\, Z_\gamma^2$ and $\mathbb{E}\, Z_\gamma^3$ are polynomials $p_2\bigl(e^{2\mathrm{i}\gamma}\bigr)$ and $p_3\bigl(e^{2\mathrm{i}\gamma}\bigr)$ in $e^{2\mathrm{i}\gamma}$ of degrees at most 2 and 3. By (6.1),

$$p_3(z)^2 = (\mathbb{E}\, V^3)^2 p_2(z)^3 \tag{6.4}$$

for every z with $|z| = 1$, and thus for every complex z.

If $\mathbb{E}\, V^3 = 0$, then $\mathbb{E}\, Z_\gamma^3 = 0$ by (6.1) and thus (ii) holds by (6.3).

Suppose now $\mathbb{E}\, V^3 \neq 0$. If $\mathbb{E}\, Z^2 = 0$, then $p_2(z)$ has degree 1 or 0 by (6.2). Degree 1 is impossible by (6.4), and thus $\mathbb{E}\,|Z|^2 = 0$ by (6.2) so $Z = 0$ a.s., and both (i) and (ii) hold.

Finally, if $\mathbb{E}\, V^3 \neq 0$ and $\mathbb{E}\, Z^2 \neq 0$, then (6.4) implies that p_2 has a double root, so its discriminant $(\mathbb{E}\,|Z|^2)^2 - \mathbb{E}\, Z^2\,\mathbb{E}\,\bar{Z}^2 = 0$, and $|\mathbb{E}\, Z^2| = \mathbb{E}\,|Z|^2$. It follows that the argument of Z^2 is constant a.s., and thus (i) holds. ☐

Returning to Theorem 2.1(iii), we can use the moments computed in Remark 2.3. Problem 6.2 shows that there really are oscillations as conjectured in Conjecture 6.1, as soon as $|\mathbb{E}(\widetilde{Z}_m - \mathbb{E}\,\widetilde{Z}_m)^2| < \mathbb{E}\,|\widetilde{Z}_m - \mathbb{E}\,\widetilde{Z}_m|^2$ and $\mathbb{E}(\widetilde{Z}_m - \mathbb{E}\,\widetilde{Z}_m)^3 \neq 0$.

It should be possible to verify this for all $m \geq 7$, perhaps using asymptotic expansions for large m, but for simplicity we have resorted to numerical verification (by `Maple`) for $m \leq 100$. We have also done the same for BST and Theorem 2.7(ii), using the moments given in Remark 2.9. We thus conclude that Conjecture 6.1 holds for RRT at least for $7 \leq m \leq 100$, and for BST at least for $9 \leq m \leq 100$.

Remark 6.3. The fact that $\mathbb{E}(\widetilde{Z}_m - \mathbb{E}\,\widetilde{Z}_m)^3 \neq 0$ also implies by (6.3) that the subsequence limit $X_\gamma := \operatorname{Re}(e^{\mathrm{i}\gamma} Z)$ has a non-zero third central moment, except for at most 6 values of $\gamma \in [0, 2\pi)$; in particular, X_γ is not normal except possibly for a few exceptional γ. Presumably, these too could be eliminated by considering fourth or fifth moments, but we have not pursued that.

References

[1] D. Aldous, The continuum random tree II: an overview. *Stochastic Analysis (Proc., Durham, 1990)*, 23–70, London Math. Soc. Lecture Note Ser. 167, Cambridge Univ. Press, Cambridge, 1991.

[2] K.B. Athreya & S. Karlin, Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39** (1968), 1801–1817.
[3] K.B. Athreya & P.E. Ney, *Branching Processes.* Springer, Berlin, 1972.
[4] S. Bernstein, Nouvelles applications des grandeurs aléatoires presqu'indépendantes. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **4** (1940), 137–150.
[5] S. Bernstein, Sur un problème du schéma des urnes à composition variable. *C. R. (Doklady) Acad. Sci. URSS (N.S.)* **28** (1940), 5–7.
[6] M. Bousquet-Mélou & S. Janson, The density of the ISE and local limit laws for embedded trees. *Appl. Probab.*,to appear August 2006.
[7] B. Chauvin, M. Drmota & J. Jabbour-Hattab, The profile of binary search trees. *Ann. Appl. Probab.* **11** (2001), no. 4, 1042–1062.
[8] B. Chauvin, T. Klein, J.-F. Marckert & A. Rouault, Martingales and profile of binary search trees. *Electron. J. Probab.* **10** (2005), 420–435.
[9] B. Chauvin & N. Pouyanne, m-ary search trees when $m \geq 27$: a strong asymptotics for the space requirements. *Random Struct. Alg.* **24** (2004), no. 2, 133–154.
[10] H.-H. Chern, M. Fuchs & H.-K. Hwang, Phase changes in random point quadtrees. *ACM Transactions on Algorithms*, to appear.
[11] H.-H. Chern & H.-K. Hwang, Phase changes in random m-ary search trees and generalized quicksort. *Random Struct. Alg.* **19** (2001), no. 3–4, 316–358.
[12] D.S. Dean & S.N. Majumdar, Phase transition in a random fragmentation problem with applications to computer science. *J. Phys. A: Math. Gen.* **35** (2002), L501–L507.
[13] L. Devroye, Branching processes and their applications in the analysis of tree structures and tree algorithms. *Probabilistic methods for algorithmic discrete mathematics*, 249–314, eds. M. Habib et al., Algorithms Combin. 16, Springer, Berlin, 1998.
[14] M. Drmota & B. Gittenberger, On the profile of random trees. *Random Struct. Alg.* **10** (1997), no. 4, 421–451.
[15] M. Drmota & H.-K. Hwang, Bimodality and phase transitions in the profile variance of random binary search trees, *SIAM J. Discrete Math.*, **19** (2005), no. 1, 19–45.
[16] M. Drmota & H.-K. Hwang, Profiles of random trees: correlation and width of random recursive trees and binary search trees. *Adv. in Appl. Probab.* **37** (2005), no. 2, 321–341.
[17] J.A. Fill & N. Kapur, Transfer theorems and asymptotic distributional results for m-ary search trees. *Random Struct. Alg.* **26** (2005), no. 4, 359–391.
[18] P. Flajolet & R. Sedgewick, *Analytic Combinatorics.* Book manuscript, August 2005. Available at `http://algo.inria.fr/flajolet/Publications/books.html`
[19] D.A. Freedman, Bernard Friedman's urn. *Ann. Math. Statist.* **36** (1965), 956–970.
[20] B. Friedman, A simple urn model. *Comm. Pure Appl. Math.* **2** (1949), 59–70.
[21] M. Fuchs, H.-K. Hwang, & R. Neininger, Profiles of random trees: limit theorems for random recursive trees and binary search trees. *Algorithmica*, to appear.
[22] A. Gut, *Probability: A Graduate Course.* Springer, New York, 2005.
[23] S. Janson, Limit theorems for certain branching random walks on compact groups and homogeneous spaces. *Ann. Probab.* **11** (1983), 909-930.
[24] S. Janson, *Gaussian Hilbert Spaces.* Cambridge Univ. Press, Cambridge, 1997.
[25] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110** (2004), no. 2, 177–245.
[26] S. Janson & R. Neininger, The size of random fragmentation trees. In preparation.
[27] H. Kesten & B.P. Stigum, Additional limit theorems for indecomposable multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37** (1966), 1463–1481.
[28] D.E. Knuth, *The Art of Computer Programming. Vol. 3: Sorting and Searching.* 2nd ed., Addison-Wesley, Reading, Mass., 1998.
[29] W. Lew & H.M. Mahmoud, The joint distribution of elastic buckets in multiway search trees. *SIAM J. Comput.* **23** (1994), no. 5, 1050–1074.

[30] G. Louchard, W. Szpankowski & J. Tang, Average profile of the generalized digital search tree and the generalized Lempel-Ziv algorithm. *SIAM J. Comput.* **28** (1999), no. 3, 904–934.

[31] H.M. Mahmoud & B. Pittel, Analysis of the space of search trees under the random insertion algorithm. *J. Algorithms* **10** (1989), no. 1, 52–75.

[32] A. Meir & J.W. Moon, On the altitude of nodes in random trees. *Canad. J. Math.* **30** (1978), 997–1015.

[33] R.T. Smythe & H. Mahmoud, A survey of recursive trees. *Theory Probab. Math. Statist.* **51** (1995), 1–27.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: `svante.janson@math.uu.se`

URL: `http://www.math.uu.se/~svante/`