WIDTH AND MODE OF THE PROFILE FOR RANDOM TREES OF LOGARITHMIC HEIGHT[1]

Luc Devroye
School of Computer Science
McGill University
Montreal, H3A 2K6
Canada

Hsien-Kuei Hwang[2]
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

August 21, 2005

Abstract

We propose a new, direct, correlation-free approach based on central moments of profiles to the asymptotics of width (size of the most abundant level) in random trees of logarithmic height. The approach is simple but gives very precise estimates for expected width, central moments of the width, and almost sure convergence. It is widely applicable to random trees of logarithmic height, including recursive trees, binary search trees, quadtrees, plane-oriented ordered trees and other varieties of increasing trees.

Abbreviated title: Width of random log trees.
MSC 2000 subject classifications: Primary 60C05; secondary 05C05 68P10.
Key words: random recursive trees, random search trees, width, profile, central moments

---

[1]This paper was prepared while both authors were visiting Institut für Stochastik und Mathematische Informatik, J. W. Goethe-Universität (Frankfurt); they thank Ralph Neininger and the the Institute for hospitality and support.
[2]Partially supported by a grant from National Science Council.

1 Introduction

Most random trees in the discrete probability literature have height either of order $\sqrt{n}$ or of order $\log n$ (n being the tree size); see Aldous (1991). For simplicity, we call these trees *square-root trees* and *log trees*, respectively. Profiles (number of nodes at each level of the tree) of random square-root trees have a rich connection to diverse structures in combinatorics and in probability and have been extensively studied. In contrast, profiles of random log trees, arising mostly from data structures and computer algorithms, were less addressed, and only quite recently were their limit behaviors, drastically different from those of square-root trees, better understood; see Drmota and Hwang (2005a, 2005b), Fuchs et al. (2005) and the references therein. We study in this paper the asymptotics of width, which is defined to be the size of the most abundant level, and its close connection to profile. Not only the results we derived are new, but also the methods of proof are of general applicability.

Recursive trees. A prototypical log tree is the recursive tree, which has been introduced in diverse fields due to its simple construction. We will present our methods of proof for recursive trees and then indicate the required elements needed for other random trees.

Combinatorially, recursive trees are rooted, labeled, non-planar trees such that the labels along any path down from any node form an increasing sequence. By *random recursive trees*, we assume that all recursive trees of n nodes are equally likely. *Probabilistically*, they can be constructed by successively adding nodes as follows. Start from a single root node with label 1. Then at the i-th stage, the new node with label i chooses any of the previous $i-1$ nodes uniformly at random (each with probability $1/(i-1)$) and is then attached to that node. This construction implies that there are $(n-1)!$ recursive trees of size n. See Drmota and Hwang (2005b) and Fuchs et al. (2005) for more references on the literature of recursive trees and their uses in other fields.

Note that the term "recursive trees" is less specific and has also been used in different contexts for different objects. For example, they are used in recursion computation theory to represent computable set of strings with branching structure, and in compiler for recording history of recursive procedures. They also appeared in classification trees, dynamic systems, and database languages with different meaning.

Profile. Let $Y_{n,k}$ denote the number of nodes at distance k from the root in random recursive trees of n nodes (the root being at level zero). Such a profile is very informative and closely related to many other shape parameters, although it does not uniquely characterize the tree. It exhibits many interesting phenomena such as (i) bimodality of the variance, (ii) different ranges for convergence in distribution and for convergence of all moments of the normalized profile $Y_{n,k}/\mathbb{E}\{Y_{n,k}\}$, ($iii$) no convergence to fixed limit law at the middle levels $k = \log n + O(1)$, and (iv) sharp sign changes for the correlation coefficients of two level-sizes; see Drmota and Hwang (2005b), Fuchs et al. (2005) for more information.

For simplicity, write throughout this paper $L_n := \max\{\log n, 1\}$. The expected profile $\mu_{n,k} := \mathbb{E}\{Y_{n,k}\}$, which gives the first picture of the general silhouette of random recursive trees, is known to be enumerated by the signless Stirling numbers of the first kind (see Fuchs et al., 2005)

$$\sum_k \mu_{n,k} u^k = \prod_{1\le j<n} \left(1+\frac{u}{j}\right) = \binom{n+u-1}{n-1}.$$

From this, it follows by saddle-point method that

$$\mu_{n,k} = \frac{n}{\sqrt{2\pi L_n}} e^{-\Delta^2/(2L_n)+O(|\Delta|^3/L_n^2)} \left(1+O\left(\frac{1+|\Delta|}{L_n}\right)\right), \tag{1}$$

uniformly for $k = L_n + O(L_n^{2/3})$, where, *here and throughout this paper*, $\Delta := k - L_n$. The asymptotic approximation (1) is crucial for our analysis. In particular, we have

$$\max_k \mu_{n,k} = \frac{n}{\sqrt{2\pi L_n}} \left(1+O(L_n^{-1})\right); \tag{2}$$

see Hwang (1995) for details and more precise expansions for $\mu_{n,k}$.

Expected width. We define the width of random recursive trees to be $W_n := \max_k Y_{n,k}$.

Theorem 1. *The expected width satisfies*

$$\mathbb{E}\{W_n\} = \frac{n}{\sqrt{2\pi L_n}} \left(1+\Theta\left(L_n^{-1}\right)\right). \tag{3}$$

This result improves upon the error term $O(L_n^{-1/4}\log L_n)$ given in Drmota and Hwang (2005b), where the proof depends on estimates for correlations of two level-sizes and tightness arguments for process. The approximation (3) also says, when compared with (1), that

$$\mathbb{E}\{W_n\} = \mu_{n,L_n+O(1)} \left(1+O(L_n^{-1})\right).$$

In particular, by (2),

$$\mathbb{E}\left\{\max_k Y_{n,k}\right\} = \max_k \mathbb{E}\{Y_{n,k}\} \left(1+O\left(L_n^{-1}\right)\right).$$

Note that the index $\hat{k}$ reaching the maximum of $\mu_{n,k}$ satisfies

$$\hat{k} = \left\lfloor L_n - 1 + \gamma + O\left(L_n^{-1}\right)\right\rfloor;$$

see pp. 140–141 of Hwang (1994) or Hammersley (1951). Erdős (1953) showed that $\hat{k}$ is unique.

An estimate for absolute central moments. We see from (3) that the expected width is asymptotically of the same order as the expected level-sizes at $k = L_n + O(1)$. We show that not only their expected values are of the same order, but also all higher absolute central moments are asymptotically close.

Theorem 2. *For any* $s \ge 0$

$$\mathbb{E}\left\{|W_n - \mathbb{E}\{W_n\}|^s\right\} = O\left(n^s L_n^{-3s/2}\right). \tag{4}$$

From Fuchs et al. (2005), we have

$$\mathbb{E}\{(Y_{n,k}-\mu_{n,k})^m\} = O\left(|\Delta|^m L_n^{-m}\mu_{n,k}^m\right) \qquad (k = L_n + o(L_n)). \tag{5}$$

By Lyapounov's inequality (see p. 174, Loève, 1977), we obtain, for any $s \geq 0$,

$$\mathbb{E}\{|Y_{n,k}-\mu_{n,k}|^s\} = O\left(|\Delta|^s L_n^{-s}\mu_{n,k}^s\right) \qquad (k = L_n + o(L_n)). \tag{6}$$

In particular, it implies, by (1), that

$$\mathbb{E}\{|W_n - \mathbb{E}\{W_n\}|^s\} = O\left(\mathbb{E}\{|Y_{n,k}-\mu_{n,k}|^s\}\right) \qquad (s \geq 0),$$

for $k = L_n + O(1)$.

Almost sure convergence. As an application of (4), we show that

$$\frac{W_n}{\mathbb{E}\{W_n\}} \longrightarrow 1 \quad \text{almost surely.} \tag{7}$$

This result was proved in Drmota and Hwang (2005b) by martingale arguments and complex analysis, following Chauvin et al. (2001). Our proof relies on (4) with $s = 2 + \varepsilon$ and the usual Borel-Cantelli argument. It is conceptually simpler and also applies to random trees for which no martingale structure is available.

Level reaching the width. Let k^* denote the level such that $Y_{n,k^*} = W_n$. To avoid ambiguity, we take k^* to be the one closest to $\lfloor L_n \rfloor$ if there are several of them. We show that k^* takes most likely the values $L_n + O(1)$.

Theorem 3. *For every $B > 0$, there exists $T_0 > 1$ such that*

$$\mathbb{P}\left(|k^* - L_n| \geq T\right) = O\left(T^{-B}\right),$$

for $T > T_0$.

Thus width will with very small probability lie outside the range $L_n + O(1)$.

Generality of the phenomena. The diverse properties we derived for the width of random recursive trees will turn out to be the tip of an iceberg. The same types of estimates will be shown, by extending the same methods of proof, to hold for a wide varieties of random trees: quadtrees, grid-trees, generalized m-ary search trees, and increasing trees. While one may expect that the same phenomena hold for general random split trees (which cover most trees we discuss as special cases; see Devroye, 1998), the main hard parts are always the uniform estimates for the expected profile for which a general uniform asymptotic tool is still lacking.

Approaches used. The most notable feature of our method of proof is that with the two crucial estimates (1) and (5) at hand, only basic probability tools such as Markov and Chebyshev inequalities and Borel-Cantelli Lemma are used. However, asymptotic tools for proving the two estimates for general random trees may differ from one case to another. In most cases we considered, the estimate (1) is proved by a combination of diverse analytic tools such as differential equations, singularity analysis (see Flajolet and Odlyzko, 1990) and saddle-point method. The remaining analysis required for higher central moments of the profile is then mostly elementary since this corresponds roughly to the large "toll-functions" cases for the underlying recurrences; see Chern et al. (2002, 2005).

Organization of the paper. For self-containedness and for paving the way for general random trees, we give a sketch of proof for (1) and (5) in the next section. We then prove the theorems in Section 3. Extension of the same arguments to other log trees is given in Sections 4–7.

Notation. Throughout this paper, the generic symbol $\varepsilon > 0$ always represents a sufficiently small constant whose value may differ from one occurrence to another. Also $L_n := \max\{\log n, 1\}$.

2 Estimates for the profile moments

We briefly sketch the main ideas leading to the estimates (1) and (5); see Fuchs et al. (2005) for details and more precise estimates than (5).

Recurrence of $Y_{n,k}$. By construction, the profile of random recursive trees satisfies the recurrence

$$Y_{n,k} \stackrel{d}{=} \sum_{1\le s<n} \frac{1}{s!} \sum_{j_1+\cdots+j_s=n-1} \underbrace{\binom{n-1}{j_1,\ldots,j_s}\frac{(j_1-1)!\cdots(j_s-1)!}{(n-1)!}}_{\mathbb{P}\left(\substack{\text{the root degree equals } s \text{ and}\\ \text{the } s \text{ subtrees have sizes } j_1,\ldots,j_s}\right)} \left(Y^{(1)}_{j_1,k-1}+\cdots+Y^{(s)}_{j_s,k-1}\right),$$

for $n \ge 2$ and $k \ge 1$ with $Y_{1,0} = 1$, where the $Y^{(i)}_{n,k}$'s are independent copies of $Y_{n,k}$. From this we deduce, by conditioning on the size of the first subtree, that

$$Y_{n,k} \stackrel{d}{=} Y_{I_n,k-1} + Y^*_{n-I_n,k} \qquad (n \ge 2; k \ge 1), \tag{8}$$

with $Y_{1,0} = 1$, where the $Y^*_{n,k}$'s are independent copies of $Y_{n,k}$ and independent of I_n, which is uniformly distributed in $\{1,\ldots,n-1\}$.

The expected profile and the expansion (1). From (8), we derive, by taking expectation and by solving the resulting recurrence, the relation

$$\sum_k \mu_{n,k} u^k = \binom{n+u-1}{n-1} \qquad (u \in \mathbb{C}).$$

Then by singularity analysis (see Flajolet and Odlyzko, 1990),

$$\sum_k \mu_{n,k} u^k = \frac{n^u}{\Gamma(1+u)}\left(1 + O\left(|u|^2 n^{-1}\right)\right), \tag{9}$$

where the O-term holds uniformly for $|u| \le C$, for any $C > 0$.

The uniform approximation (1) is then obtained by Cauchy's integral formula using (9) and the saddle-point method; see Hwang (1995).

A uniform estimate for $\mu_{n,k}$. A very useful uniform estimate for $\mu_{n,k}$ is given by

$$\mu_{n,k} = O\left(L_n^{-1/2} r^{-k} n^r\right) \qquad (0 < r = O(1)), \tag{10}$$

uniformly for all $0 \le k \le n$. This is easily obtained by Cauchy's integral formula and (9) since

$$\mu_{n,k} = O\left(r^{-k} n^r \int_{-\pi}^{\pi} n^{-r(1-\cos t)} \mathrm{d}t\right),$$

which gives (10). *Throughout this paper*, r is always taken to be $r = 1 + o(1)$ unless otherwise specified.

Although one can prove that $\mu_{n,k} = O(L_n^k/k!)$ for $0 \le k \le n$, the reason of using the estimate (10) instead of $O(L_n^k/k!)$ for all k is that for general random search trees it is much harder to derive the Poisson type estimate for all k.

Recurrence of higher central moments. Let $P_{n,k}^{(m)} = \mathbb{E}\{(Y_{n,k} - \mu_{n,k})^m\}$. Then $P_{n,k}^{(m)}$ satisfies the recurrence

$$P_{n,k}^{(m)} = \frac{1}{n-1} \sum_{1 \le j < n} \left(P_{j,k-1}^{(m)} + P_{n-j,k}^{(m)}\right) + Q_{n,k}^{(m)},$$

with $P_{n,0}^{(m)} = 0$ for $n, m \ge 1$, where

$$Q_{n,k}^{(m)} := \sum_{(a,b,c)\in\mathcal{I}_m} \binom{m}{a,b,c} \frac{1}{n-1} \sum_{1\le j<n} P_{j,k-1}^{(a)} P_{n-j,k}^{(b)} \nabla_{n,k}^c(j) \qquad (m \ge 2),$$

with $\nabla_{n,k}(j) := \mu_{j,k-1} + \mu_{n-j,k} - \mu_{n,k}$ and

$$\mathcal{I}_m := \{(a,b,c) \in \mathbb{Z}^3 \,:\, a+b+c = m, 0 \le a, b < m, 0 \le c \le m\}.$$

We prove (5) in two stages. A uniform estimate for $\nabla_{n,k}(j)$ for $1 \le j, k < n$ is first derived, which then implies by induction a uniform bound for $P_{n,k}^{(m)}$ for $1 \le k < n$. This bound is however not tight when $\Delta = o(\sqrt{L_n})$. Then we refine the estimate for $\nabla_{n,k}(j)$ when $\Delta = O(\sqrt{L_n})$, which then leads to (5) by another induction.

First estimate for $P_{n,k}^{(m)}$. By (9), we have the integral representation

$$\nabla_{n,k}(j) = \frac{1}{2\pi i} \oint_{|u|=r} \frac{u^{-k-1} n^u}{\Gamma(1+u)} \varphi(u; j/n) \left(1 + O\left(j^{-1} + (n-j)^{-1}\right)\right) \mathrm{d}u, \tag{11}$$

where $\varphi(u; x) := u x^u + (1-x)^u - 1$. Since $\varphi(1; x) = 0$, we have

$$\frac{\varphi(u;x)}{\Gamma(1+u)} = O\left(|u-1|\right),$$

uniformly for $x \in [0, 1]$. Substituting this estimate in (11), we obtain

$$\begin{aligned}\nabla_{n,k}(j) &= O\left(r^{-k} n^r \int_{-\pi}^{\pi} \left|re^{i\theta} - 1\right| n^{-r(1-\cos\theta)} \mathrm{d}\theta\right) \\ &= O\left((|r-1| + L_n^{-1/2}) L_n^{-1/2} r^{-k} n^r\right),\end{aligned} \tag{12}$$

uniformly for $1 \le j, k < n$, where $r = 1 + o(1)$. This bound is not tight for all k but is sufficient for most of our purposes. In particular, since r is not specially chosen to minimize the error term, (12) is not optimal when $|r-1| = o(L_n^{-1/2})$, which is the case when we choose $r = k/L_n$ and $|k - L_n| = o(\sqrt{L_n})$.

We now prove by induction that

$$P_{n,k}^{(m)} = O\left(\left(|r-1|^m + L_n^{-m/2}\right) L_n^{-m/2} r^{-km} n^{mr}\right) \qquad (m \ge 0), \tag{13}$$

uniformly for $1 \le k < n$.

Obviously, (13) holds for $m = 0, 1$. Assume $m \ge 2$. To estimate $Q_{n,k}^{(m)}$, we split the sum into two parts.

$$Q_{n,k}^{(m)} = \sum_{(a,b,c)\in\mathcal{I}_m} \binom{m}{a,b,c} \frac{1}{n-1} \left(\sum_{j\in\mathcal{J}_m} + \sum_{j\in\mathcal{J}_m'}\right) P_{j,k-1}^{(a)} P_{n-j,k}^{(b)} \nabla_{n,k}^{c}(j),$$

where $\mathcal{J}_m := \{j \ : \ n/L_n^m \le j \le n - n/L_n^m\}$ and $\mathcal{J}_m' := \{1, \dots, n-1\} \setminus \mathcal{J}_m$.

Then by induction and (12), the terms in $Q_{n,k}^{(m)}$ with $j \in \mathcal{J}_m'$ are bounded above by

$$\begin{aligned} O\Bigg(r^{-mk} n^{-1} \sum_{(a,b,c)\in\mathcal{I}_m} \bigg(& L_n^{-(b+c)/2} n^{(b+c)r} \sum_{j<n/L_n^m} L_j^{-a/2} j^{ar} \\ & + L_n^{-(a+c)/2} n^{(a+c)r} \sum_{j<n/L_n^m} L_j^{-b/2} j^{br} \bigg)\Bigg) \\ = O\left(L_n^{-3m/2} r^{-mk} n^{mr}\right), & \end{aligned}$$

uniformly for $1 \le k < n$.

On the other hand, when $j \in \mathcal{J}_m$, we have $L_j \sim L_{n-j} \sim L_n$; thus by induction and the two estimates (12) and (13)

$$\begin{aligned} &\sum_{(a,b,c)\in\mathcal{I}_m} \binom{m}{a,b,c} \frac{1}{n-1} \sum_{j\in\mathcal{J}_m} P_{j,k-1}^{(a)} P_{n-j,k}^{(b)} \nabla_{n,k}^{c}(j) \\ &\qquad = O\left(\left(|r-1|^m + L_n^{-m/2}\right) L_n^{-m/2} r^{-km} n^{mr}\right), \end{aligned}$$

it follows that

$$Q_{n,k}^{(m)} = O\left(\left(|r-1|^m + L_n^{-m/2}\right) L_n^{-m/2} r^{-km} n^{mr}\right), \tag{14}$$

uniformly for $1 \le k < n$.

From Fuchs et al. (2005), we have the closed-form expression

$$P_{n,k}^{(m)} = Q_{n,k}^{(m)} + \sum_{1\le j<n} \sum_{0\le \ell\le k} \frac{Q_{j,k-\ell}^{(m)}}{j} [u^\ell](u+1) \prod_{j<h<n} \left(1 + \frac{u}{h}\right), \tag{15}$$

where $[u^\ell]F(u)$ denotes the coefficient of u^ℓ in the Taylor expansion of F. Substituting the estimate (14), we obtain

$$P_{n,k}^{(m)} = O\Bigg(Q_{n,k}^{(m)} + r^{-km} \sum_{1\le j<n} \left(|r-1|^m + L_j^{-m/2}\right) L_j^{-m/2} j^{mr-1} \\ \times \sum_{0\le \ell\le k} r^{m\ell}[u^\ell](u+1) \prod_{j<h<n}\left(1+\frac{u}{h}\right)\Bigg).$$

Now

$$\sum_{0\le\ell\le k} r^{m\ell}[u^\ell](u+1)\prod_{j<h<n}\left(1+\frac{u}{h}\right) \le (1+r^m)\prod_{j<h<n}\left(1+\frac{r^m}{h}\right) \\ = O\left(\left(\frac{n}{j}\right)^{r^m}\right).$$

Thus (13) follows.

When $k \sim L_n$, we take $r = k/L_n$ in (13), giving

$$P_{n,k}^{(m)} = O\left(\left(|\Delta|^m + L_n^{m/2}\right) L_n^{-m}\mu_{n,k}^m\right),$$

which proves (5) when $\sqrt{L_n} \le |\Delta| = o(L_n)$.

Proof of (5) when $\Delta = O(\sqrt{L_n})$. We now refine the above procedure and prove (5) when $\Delta = O(\sqrt{L_n})$, which has the form

$$P_{n,k}^{(m)} = O\left(|\Delta|^m L_n^{-3m/2} n^m\right) \qquad (m \ge 0). \tag{16}$$

By applying the expansion

$$\varphi(u;x) = \varphi_u'(1;x)(u-1) + O\left(|u-1|^2\right) \qquad (x\in[0,1]),$$

and the usual saddle-point method to (11), we deduce that

$$\nabla_{n,k}(j) = O\left(|\Delta| L_n^{-3/2} n\right), \tag{17}$$

uniformly for $\Delta = O(\sqrt{L_n})$ and $1 \le j < n$. Note that this estimate also follows from (1).

By the same procedure used to prove (14) and by applying (13) to terms with $j \in \mathcal{J}_m'$, we have

$$Q_{n,k}^{(m)} = O\left(|\Delta|^m L_n^{-3m/2} n^m\right) \qquad (m \ge 2),$$

uniformly for $\Delta = O(\sqrt{L_n})$. This estimate and (14) gives, by (15) and a similar decomposition of the sums involved,

$$P_{n,k}^{(m)} = O\Bigg(Q_{n,k}^{(m)} + \sum_{j\in\mathcal{J}_m} \sum_{0\le \ell = o(L_n)} |k-\ell-L_j|^m L_j^{-3m/2} j^{m-1}[u^\ell](u+1)\prod_{j<h<n}\left(1+\frac{u}{h}\right)\Bigg) \\ = O\left(|\Delta|^m L_n^{-3m/2} n^m\right).$$

This proves (16). $\square$

Such a two-stage proof of (5) is completely general when we have an integral representation for $\nabla_{n,k}(j)$ of the form (11) and a closed-form similar to (15). We will propose means of handling the cases when no closed-form solution like (15) is available.

3 Asymptotics of the moments of the width

We first prove Theorem 1; then we extend the proof for (4) and finally prove Theorem 3.

3.1 Expected width

Lower bound for the expected width. The lower bound follows easily from the inequality

$$\mathbb{E}\{W_n\} \geq M_n,$$

where

$$M_n := \max_k \mathbb{E}\{Y_{n,k}\} = \frac{n}{\sqrt{2\pi L_n}}\left(1 + O\left(L_n^{-1}\right)\right);$$

see (2).

An inequality for the upper bound. For the upper bound, we use the inequality

$$\begin{aligned}\mathbb{E}\{W_n\} &\leq M_n + \sum_{|\Delta|\leq K} \mathbb{E}\left\{(Y_{n,k} - M_n)_+\right\} + \sum_{|\Delta|>K} \mu_{n,k} \\ &=: w_n^{(1)} + w_n^{(2)} + w_n^{(3)},\end{aligned} \tag{18}$$

where $K := L_n^{2/3}$.

The sum $w_n^{(3)}$. The last sum is easily estimated since by (10)

$$w_n^{(3)} = O\left(L_n^{-1/2} n^r \left(\sum_{0\leq k\leq L_n - K} + \sum_{k\geq L_n+K}\right) r^{-k}\right).$$

Taking $r = 1 - L_n^{-1/3}$, we see that

$$\begin{aligned}\sum_{0\leq k\leq L_n-K} \mu_{n,k} &= O\left(L_n^{-1/2} n^r \frac{r^{-L_n+K}}{1-r}\right) \\ &= O\left(L_n^{-1/2} n^{1-L_n^{-1/3}} L_n^{1/3} \left(1 - L_n^{-1/3}\right)^{-L_n + L_n^{2/3}}\right) \\ &= O\left(n L_n^{-1/6} e^{-L_n^{1/3}/2}\right),\end{aligned}$$

and the same upper bound holds for $\sum_{k\geq L_n+K} \mu_{n,k}$ by taking $r = 1 + L_n^{-1/3}$.

An estimate for the second sum $w_n^{(2)}$. We use the inequalities

$$\begin{aligned}\mathbb{E}\{(Y_{n,k}-M_n)_+\} &\le \mathbb{E}\left\{(Y_{n,k}-\mu_{n,k})\mathbf{1}_{(Y_{n,k}>M_n)}\right\}\\ &\le \frac{\mathbb{E}\{(Y_{n,k}-\mu_{n,k})^2\}}{M_n-\mu_{n,k}},\end{aligned}$$

for those k's for which $M_n > \mu_{n,k}$. By (1)

$$\begin{aligned}M_n-\mu_{n,k} &= \frac{n}{\sqrt{2\pi L_n}}\left(1-e^{-\Delta^2/(2L_n)+O(|\Delta|^3/L_n^2)}\right)(1+o(1))\\ &\ge \frac{n}{2\sqrt{2\pi L_n}}\left(1-e^{-\Delta^2/(3L_n)}\right)(1+o(1)),\end{aligned} \tag{19}$$

uniformly for $1 \le |\Delta| \le K$. On the other hand, we also have the estimates for the variance

$$\mathbb{V}\{Y_{n,k}\} = O\left(\Delta^2 L_n^{-2}\mu_{n,k}^2\right) = O\left(\Delta^2 L_n^{-3} n^2 e^{-\Delta^2/L_n}\right), \tag{20}$$

uniformly for $1 \le |\Delta| \le K$. It follows from these estimates that

$$\begin{aligned}w_n^{(2)} &\le \sqrt{\mathbb{V}\{Y_{n,\lfloor L_n\rfloor}\}} + \sum_{1\le|\Delta|\le K}\frac{\mathbb{V}\{Y_{n,k}\}}{M_n-\mu_{n,k}}\\ &= O\left(nL_n^{-3/2}\right) + O\left(L_n^{-5/2} n\int_1^\infty \frac{x^2 e^{-x^2/L_n}}{1-e^{-x^2/(3L_n)}}\,\mathrm{d}x\right)\\ &= O\left(nL_n^{-1}\right).\end{aligned}$$

Collecting all estimates, we get a weaker error term than (3)

$$\mathbb{E}\{W_n\} = \frac{n}{\sqrt{2\pi L_n}}\left(1+O\left(L_n^{-1/2}\right)\right), \tag{21}$$

but we only used estimates for $\mathbb{E}\{Y_{n,k}\}$ and $\mathbb{V}\{Y_{n,k}\}$.

Improving the error term by fourth central moments of $Y_{n,k}$. We can improve the error term in (21) by using the estimate (5) for the fourth central moment of $Y_{n,k}$. Taking $m = 4$ in (5) and repeating the same analysis as above

$$\begin{aligned}w_n^{(2)} &\le \sqrt{\mathbb{V}\{Y_{n,\lfloor L_n\rfloor}\}} + \sum_{1\le|\Delta|\le K}\frac{\mathbb{E}\left\{(Y_{n,k}-\mu_{n,k})^4\right\}}{(M_n-\mu_{n,k})^3}\\ &= O\left(nL_n^{-3/2}\right) + O\left(nL_n^{-9/2}\int_1^\infty \frac{x^4 e^{-2x^2/L_n}}{(1-e^{-x^2/(3L_n)})^3}\,\mathrm{d}x\right)\\ &= O\left(nL_n^{-3/2}\right) + O\left(nL_n^{-2}\int_{1/L_n}^\infty \frac{v^{3/2}e^{-2v}}{(1-e^{-v/3})^3}\,\mathrm{d}v\right)\\ &= O\left(nL_n^{-3/2}\right) + O\left(nL_n^{-2}\int_{1/L_n}^\infty v^{-3/2}\mathrm{d}v\right)\\ &= O\left(nL_n^{-3/2}\right).\end{aligned}$$

This proves (3).

3.2 Higher absolute central moments of W_n

We prove only an upper bound for $s = 2$, namely for the variance of W_n, other values of s following by the same argument and Lyapounov's inequality.

An upper bound for the variance of the width. We show, by using central moments of $Y_{n,k}$ of order 6, that

$$\mathbb{V}\{W_n\} = O\left(n^2 L_n^{-3}\right), \tag{22}$$

which proves (4) with $s = 2$.

The proof extends that for $\mathbb{E}\{W_n\}$. Define $k_0 = \lfloor L_n \rfloor$. We start from

$$\begin{aligned}\mathbb{E}\left\{(W_n - \mathbb{E}\{W_n\})^2\right\} &= \mathbb{E}\left\{(W_n - \mu_{n,k_0} + \mu_{n,k_0} - \mathbb{E}\{W_n\})^2\right\} \\ &\leq 2\mathbb{E}\left\{(W_n - \mu_{n,k_0})^2\right\} + 2\mathbb{E}\left\{(\mu_{n,k_0} - \mathbb{E}\{W_n\})^2\right\}.\end{aligned}$$

By (3),

$$\mathbb{E}\left\{(\mu_{n,k_0} - \mathbb{E}\{W_n\})^2\right\} = O\left(n^2 L_n^{-3}\right).$$

And, similar to the analysis for $\mathbb{E}\{W_n\}$,

$$\begin{aligned}\mathbb{E}\left\{(W_n - \mu_{n,k_0})^2\right\} &\leq \mathbb{E}\left\{\sum_{|\Delta|\geq 0} (Y_{n,k_0+\Delta} - \mu_{n,k_0})_+^2 \cdot \mathbf{1}_{(Y_{n,k_0+\Delta} > \mu_{n,k_0})}\right\} \\ &\leq \mathbb{V}\{Y_{n,k_0}\} + \sum_{1\leq|\Delta|\leq K} \mathbb{E}\left\{(Y_{n,k_0+\Delta} - \mu_{n,k_0})_+^2\right\} + \sum_{|\Delta|\geq K} \mathbb{V}\{Y_{n,k_0+\Delta}\} \\ &=: v_n^{(1)} + v_n^{(2)} + v_n^{(3)}.\end{aligned}$$

By (20),

$$v_n^{(1)} = O(n^2 L_n^{-3}).$$

The estimate for $v_n^{(2)}$ follows *mutatis mutandis* from that for $w_n^{(2)}$ by using (5) with $m = 6$.

$$\begin{aligned}v_n^{(2)} &\leq \sum_{1\leq|\Delta|\leq K} \frac{\mathbb{E}\left\{(Y_{n,k} - \mu_{n,k})^6\right\}}{(\mu_{n,k_0} - \mu_{n,k_0+\Delta})^4} \\ &= O\left(n^2 L_n^{-7} \sum_{1\leq|\Delta|\leq K} \frac{\Delta^6 e^{-3\Delta^2/L_n}}{(1 - e^{-\Delta^2/(3L_n)})^4}\right) \\ &= O\left(n^2 L_n^{-7/2} \int_{1/L_n}^{\infty} v^{-3/2} \mathrm{d}v\right) \\ &= O\left(n^2 L_n^{-3}\right).\end{aligned}$$

For the last term $v_n^{(3)}$, we use again (10)

$$\mathbb{V}\{Y_{n,k}\} \leq \mu_{n,k}^2 = O\left(L_n^{-1} r^{-2k} n^{2r}\right),$$

uniformly for $1 \leq k \leq n$, where $r > 0$ is any bounded real number. Substituting this into $v_n^{(3)}$ gives

$$\begin{aligned} v_n^{(3)} &= O\left(L_n^{-1} n^{2r} \sum_{|\Delta| \geq K} r^{-2k_0 - 2\Delta}\right) \\ &= O\left(L_n^{2/3} n^2 e^{-L_n^{1/3}}\right), \end{aligned}$$

by taking $r = 1 + \mathrm{sign}(\Delta) L_n^{-1/3}$.

This completes the proof of (22).

Higher central moments of W_n. The same analysis can be carried out for higher absolute central moments using (6). Then the same proof for $\mathbb{V}\{W_n\}$ gives (4) by using (6) with order $(2s+2)$.

Almost sure convergence. We need first a tail bound for the width. By Markov inequality (see p. 160, Loève, 1977; sometimes referred to as Chebyshev inequality)

$$\begin{aligned} \mathbb{P}\{|W_n - \mathbb{E}\{W_n\}| \geq \varepsilon \mathbb{E}\{W_n\}\} &\leq \frac{\mathbb{E}\{|W_n - \mathbb{E}\{W_n\}|^s\}}{(\varepsilon \mathbb{E}\{W_n\})^s} \\ &= O\left(\varepsilon^{-s} L_n^{-s}\right), \end{aligned}$$

for any $s > 0$ and $\varepsilon \in (0, 1)$.

From this estimate, it follows, by applying Borel-Cantelli and by taking $s > 2$, that

$$\frac{W_{n_\ell}}{\mathbb{E}\{W_{n_\ell}\}} \longrightarrow 1 \quad \text{almost surely,}$$

where $n_\ell := \lfloor e^{\sqrt{\ell}} \rfloor$, since $\sum_\ell L_{n_\ell}^{-s} = O\left(\sum_\ell \ell^{-s/2}\right) = O(1)$.

Now observe that

$$n_{\ell+1} - n_\ell = \Theta\left(n_\ell \ell^{-1/2}\right) = \Theta\left(n_\ell L_{n_\ell}^{-1}\right) = \Theta\left(\mathbb{E}\left\{W_{n_\ell}\right\} L_{n_\ell}^{-1/2}\right).$$

On the other hand, by construction, adding a new node to random recursive trees affects the value of W_n by at most 1. Consequently,

$$\sup_{n_\ell \leq n < n_{\ell+1}} \max\left(|W_n - W_{n_\ell}|, |\mathbb{E}\{W_n\} - \mathbb{E}\left\{W_{n_\ell}\right\}|\right) \leq n_{\ell+1} - n_\ell = \Theta\left(\mathbb{E}\left\{W_{n_\ell}\right\} L_{n_\ell}^{-1/2}\right).$$

So, deterministically,

$$\begin{aligned} \sup_{n_\ell \leq n < n_{\ell+1}} \left| \frac{W_n}{\mathbb{E}\{W_n\}} - \frac{W_{n_\ell}}{\mathbb{E}\left\{W_{n_\ell}\right\}} \right| &= O\left(\frac{\mathbb{E}\left\{W_{n_\ell}\right\} L_{n_\ell}^{-1/2}}{\mathbb{E}\left\{W_{n_\ell}\right\} - (n_{\ell+1} - n_\ell)}\right) \\ &= O\left(L_{n_\ell}^{-1/2}\right) \\ &= O\left(\ell^{-1/4}\right). \end{aligned}$$

This completes the proof of (7).

An alternative form to (7). The same argument can be modified to show that

$$\frac{W_n}{n/\sqrt{2\pi L_n}} = 1 + O\left(L_n^{-1+\delta}\right), \tag{23}$$

almost surely, for any fixed $\delta > 0$. The proof is modified from that for (7) as follows. By (3), we have

$$\frac{W_n}{n/\sqrt{2\pi L_n}} = \frac{W_n}{\mathbb{E}\{W_n\}}\left(1 + O\left(L_n^{-1}\right)\right).$$

Instead of $n_\ell := \lfloor e^{\sqrt{\ell}} \rfloor$, we now take $n_\ell := \lfloor e^{\sqrt{\ell}/(2-\delta)} \rfloor$. Then, setting $\varepsilon = \varepsilon_n = L_n^{-1+\delta}$ in the proof, we deduce that, again by Borel-Cantelli,

$$\frac{W_{n_\ell}}{\mathbb{E}\{W_{n_\ell}\}} = 1 + O(\varepsilon_{n_\ell}),$$

almost surely as $\ell \to \infty$ provided that $\varepsilon_{n_\ell}^{-s} L_{n_\ell}^{-s}$ is summable in ℓ. This forces the choice $s > 2/\delta$. Next,

$$n_{\ell+1} - n_\ell = \Theta\left(\mathbb{E}\{W_{n_\ell}\}\ell^{-1/4}\right).$$

This proves (23).

Almost sure convergence for $Y_{n,k}$. We can also obtain strong convergence by the same argument for the profiles $Y_{n,k}$ in the central range $[L_n - L_n^{1-\varepsilon}, L_n + L_n^{1-\varepsilon}]$, where $\varepsilon \in (0,1)$. We now prove that

$$\sup_{L_n - L_n^{1-\varepsilon} \le \kappa \le L_n + L_n^{1-\varepsilon}} \left| \frac{Y_{n,\kappa}}{\mathbb{E}\{Y_{n,\kappa}\}} - 1 \right| \to 0, \tag{24}$$

almost surely.

Proof. Set $t_n := 2L_n^{1-\varepsilon}$ and $n_\ell := \lfloor e^{\sqrt{\ell}} \rfloor$. Using (6) and Markov's inequality used above, it is easy to see that

$$\sup_{L_{n_\ell} - t_{n_\ell} \le \kappa \le L_{n_\ell} + t_{n_\ell}} \left| \frac{Y_{n_\ell,\kappa}}{\mathbb{E}\{Y_{n_\ell,\kappa}\}} - 1 \right| \to 0,$$

almost surely as $\ell \to \infty$. By the union bound and Borel-Cantelli, this requires that we take s so large that $L_{n_\ell}^{-s} t_{n_\ell}^{1+s}$ is summable in ℓ. Any choice with $s > 3/\varepsilon - 1$ suffices for that purpose. Furthermore, by the monotonicity of $Y_{n,k}$ in n for fixed k,

$$\sup_{n_\ell \le n < n_{\ell+1}} \sup_{L_{n_\ell} - t_{n_\ell} \le \kappa \le L_{n_\ell} + t_{n_\ell}} |Y_{n,\kappa} - Y_{n_\ell,\kappa}| \le \sup_{L_{n_\ell} - t_{n_\ell} \le \kappa \le L_{n_\ell} + t_{n_\ell}} |Y_{n_{\ell+1},\kappa} - Y_{n_\ell,\kappa}|,$$

and

$$\sup_{n_\ell \le n < n_{\ell+1}} \sup_{L_{n_\ell} - t_{n_\ell} \le \kappa \le L_{n_\ell} + t_{n_\ell}} |\mathbb{E}\{Y_{n,\kappa}\} - \mathbb{E}\{Y_{n_\ell,\kappa}\}| \le \sup_{L_{n_\ell} - t_{n_\ell} \le \kappa \le L_{n_\ell} + t_{n_\ell}} |\mathbb{E}\{Y_{n_{\ell+1},\kappa}\} - \mathbb{E}\{Y_{n_\ell,\kappa}\}|.$$

Thus,

$$\begin{aligned}\sup_{n_\ell \le n < n_{\ell+1}} \left| \frac{Y_{n,\kappa}}{\mathbb{E}\{Y_{n,\kappa}\}} - \frac{Y_{n_\ell,\kappa}}{\mathbb{E}\{Y_{n_\ell,\kappa}\}} \right| &\le \frac{Y_{n_{\ell+1},\kappa}}{\mathbb{E}\{Y_{n_\ell,\kappa}\}} - \frac{Y_{n_\ell,\kappa}}{\mathbb{E}\{Y_{n_\ell,\kappa}\}} \\ &\le \frac{Y_{n_{\ell+1},\kappa}}{\mathbb{E}\{Y_{n_{\ell+1},\kappa}\}} - \frac{Y_{n_\ell,\kappa}}{\mathbb{E}\{Y_{n_\ell,\kappa}\}} + \left(\frac{\mathbb{E}\{Y_{n_{\ell+1},\kappa}\}}{\mathbb{E}\{Y_{n_\ell,\kappa}\}} - 1\right)\frac{Y_{n_{\ell+1},\kappa}}{\mathbb{E}\{Y_{n_{\ell+1},\kappa}\}}.\end{aligned}$$

Putting the supremum over $L_{n_\ell} - t_{n_\ell} \le \kappa \le L_{n_\ell} + t_{n_\ell}$ in front of all of the latter inequalities, we see that both terms tend to zero almost surely provided that

$$\lim_{\ell\to\infty} \sup_{L_{n_\ell}-t_{n_\ell}\le\kappa\le L_{n_\ell}+t_{n_\ell}} \left| \frac{\mathbb{E}\{Y_{n_{\ell+1},\kappa}\}}{\mathbb{E}\{Y_{n_\ell,\kappa}\}} - 1 \right| = 0.$$

This follows from an extension of the Taylor series estimate used in (1); indeed, the estimate (see Hwang, 1995)

$$\mu_{n,k} = \frac{L_n^k}{\Gamma(1+k/L_n)k!} \left(1 + O\left(L_n^{-1}\right)\right) \qquad (k = O(L_n)),$$

is sufficient for our use.

Thus we have shown that

$$\sup_{n_\ell\le n<n_{\ell+1}} \sup_{L_{n_\ell}-t_{n_\ell}\le\kappa\le L_{n_\ell}+t_{n_\ell}} \left| \frac{Y_{n,\kappa}}{\mathbb{E}\{Y_{n,\kappa}\}} - 1 \right| \to 0$$

almost surely. An additional argument shows that for ℓ large enough, $[L_n - L_n^{1-\varepsilon}, L_n + L_n^{1-\varepsilon}]$ is contained in $[L_{n_\ell} - t_{n_\ell}, L_{n_\ell} + t_{n_\ell}]$ for $n_\ell \le n < n_{\ell+1}$, thus concluding the proof of (24).

3.3 Level reaching the width

We now prove Theorem 3. For $|\Delta| > |\hat{k} - k_0|$ and $B > 1$

$$\begin{aligned}
\mathbb{P}(k^* = k_0 + \Delta) &= \mathbb{P}(W_n = Y_{n,k_0+\Delta}) \\
&\le \mathbb{P}(Y_{n,k_0+\Delta} > Y_{n,k_0}) \\
&= \mathbb{P}\left(Y_{n,k_0+\Delta} - \mu_{n,k_0+\Delta} > Y_{n,k_0} - \mu_{n,k_0} + \mu_{n,k_0} - \mu_{n,k_0+\Delta}\right) \\
&\le \mathbb{P}\left(Y_{n,k_0+\Delta} - \mu_{n,k_0+\Delta} \ge \frac{1}{2}\left(\mu_{n,k_0} - \mu_{n,k_0+\Delta}\right)\right) \\
&\qquad + \mathbb{P}\left(Y_{n,k_0} - \mu_{n,k_0} \le -\frac{1}{2}\left(\mu_{n,k_0} - \mu_{n,k_0+\Delta}\right)\right) \\
&\le \frac{2^B\mathbb{E}|Y_{n,k_0+\Delta} - \mu_{n,k_0+\Delta}|^B}{(\mu_{n,k_0} - \mu_{n,k_0+\Delta})^B} + \frac{2^B\mathbb{E}|Y_{n,k_0} - \mu_{n,k_0}|^B}{(\mu_{n,k_0} - \mu_{n,k_0+\Delta})^B},
\end{aligned}$$

by Markov inequality. By (1), we obtain a similar estimate to (19) for $\mu_{n,k_0} - \mu_{n,k_0+\Delta}$, which together with (20) gives

$$\begin{aligned}
\mathbb{P}(k^* = k_0 + \Delta) &= O\left(\frac{\Delta^B L_n^{-B} e^{-B\Delta^2/(2L_n)}}{(1 - e^{-\Delta^2/(3L_n)})^B} + \frac{L_n^{-B}}{(1 - e^{-\Delta^2/(3L_n)})^B}\right) \\
&= O\left(\Delta^{-B} + \Delta^{-2B} \vee L_n^{-B}\right) \\
&= O\left(\Delta^{-B}\right),
\end{aligned}$$

uniformly for $1 \le |\Delta| \le K$. It follows that there exists a $T_0 > 1$ such that for $T > T_0$

$$\begin{aligned}
\mathbb{P}(|k^* - k_0| \ge T) &= O\left(\sum_{T\le|\Delta|\le K} \Delta^{-B}\right) + \mathbb{P}(|k^* - k_0| \ge K) \\
&= O\left(T^{1-B}\right) + \mathbb{P}(|k^* - k_0| \ge K).
\end{aligned}$$

The tail probability $\mathbb{P}(|k^* - k_0| \geq K)$ is estimated as follows. Let $k_1 := \lfloor \sqrt{L_n} \rfloor$.

$$\begin{aligned}
\mathbb{P}(|k^* - k_0| \geq K) &\leq \mathbb{P}\left(\max_{|k-k_0|\geq K} Y_{n,k} \geq Y_{n,k_0} \right) \\
&\leq \mathbb{P}\left(\max_{|k-k_0|\geq K} Y_{n,k} \geq \mu_{n,k_0+k_1} \right) + \mathbb{P}\left(Y_{n,k_0} < \mu_{n,k_0+k_1}\right) \\
&\leq \mu_{n,k_0+k_1}^{-1} \sum_{|k-k_0|\geq K} \mu_{n,k} + \frac{\mathbb{V}\{Y_{n,k_0}\}}{(\mu_{n,k_0} - \mu_{n,k_0+k_1})^2} \\
&= O\left(L_n^{1/3} e^{-L_n^{1/3}/2} + L_n^{-2} \right),
\end{aligned}$$

which tends to zero as $n \to \infty$, where we used again (10) to bound $\sum_{|k-k_0|\geq K} \mu_{n,k}$. Since $B > 1$ is arbitrary, this proves Theorem 3.

Limit distribution of W_n? It is known that the centered and normalized random variables $(Y_{n,k} - \mu_{n,k})/\sqrt{\mathbb{V}\{Y_{n,k}\}}$ do not converge to a fixed limit law when $k = L_n + O(1)$ due to periodicity; see Fuchs et al. (2005). The origin of the periodicity lies at the second-order term in the asymptotic expansion of $\mu_{n,L_n+O(1)}$

$$\mu_{n,k_0+\ell} = \frac{n}{\sqrt{2\pi L_n}} \left(1 + \frac{p_1(\{L_n\})}{L_n} + O\left(L_n^{-2}\right) \right) \qquad (\ell \in \mathbb{Z}),$$

where $\{x\}$ denotes the fractional part of x and

$$p_1(x) = -\frac{x^2}{2} + \left(\frac{3}{2} - \gamma + \ell\right) x - \frac{\ell^2}{2} - \left(\frac{3}{2} - \gamma\right) \ell - \frac{\gamma^2}{2} + \gamma + \frac{\pi^2}{12} - \frac{13}{12}.$$

This periodic second-order term is the origin of all fluctuations of higher central moments.

The main open question is the limit distribution (if it exists) of W_n. Simulations seem to indicate the closeness of the histogram of W_n and that of $Y_{n,\lfloor L_n \rfloor -1}$; see Figure 1.

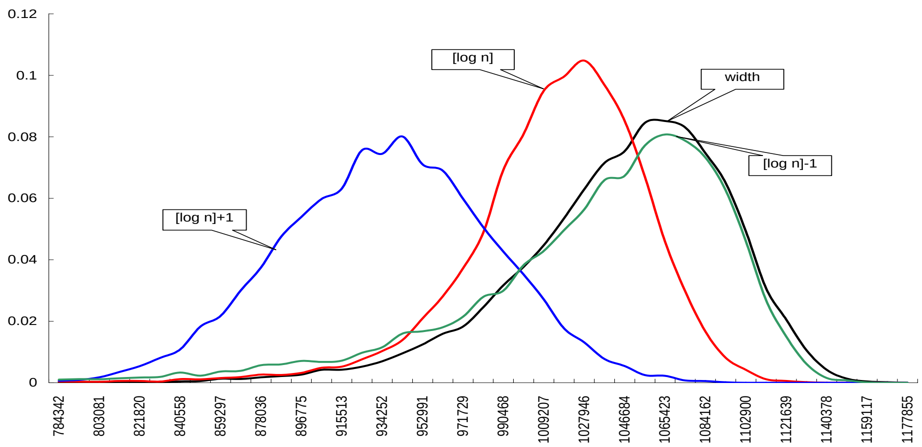

Figure 1: *Simulated histograms for W_n and $Y_{n,\lfloor L_n \rfloor + \ell}$ for $\ell = -1, \ldots, 1$, where* $n = 10^7$.

4 Width of general random log trees

We first describe roughly the estimates we need for handling the width of general random log trees, and then discuss a few concrete examples. We start from a general framework for the profile of random log trees.

Recurrence of profile. Assume that the profile of the random log trees in question satisfies

$$Y_{n,k} \stackrel{d}{=} \sum_{1\le j\le h} Y^{(j)}_{I_{n,j},k-1} \qquad (n \ge 2; k \ge 1), \tag{25}$$

with $Y_{n,0} = 1$ for $n \ge 1$, where the $Y^{(j)}_{n,k}$'s are independent copies of $Y_{n,k}$ and the underlying splitting distribution satisfies $\sum_j I_{n,j} = n - \kappa$ for some integer $\kappa \ge 0$. Then the moments of $Y_{n,k}$ satisfy the recurrence

$$a_{n,k} = h \sum_{0\le j<n} \pi_{n,j} a_{j,k-1} + b_{n,k}, \tag{26}$$

where $h \ge 2$, $\sum_j \pi_{n,j} = 1$ and $\pi_{n,j} = \mathbb{P}(I_{n,1} = j)$. For our purpose, we can always assume that $b_{n,k} = 0$ for $k < 0$ and $k \ge n$.

An analytic scheme for the expected profile. The following simple framework gives sufficient conditions we need in order to obtain asymptotics of the width of general random log trees.

Assume that the generating polynomial of the expected profile $\mu_{n,k} := \mathbb{E}\{Y_{n,k}\}$ satisfies

$$\Xi_n(u) := \sum_k \mu_{n,k} u^k = g(u) n^{f(u)} \left(1 + O\left(n^{-\varepsilon}\right)\right), \tag{27}$$

uniformly for $|u-1| \le \varepsilon_0$, $\varepsilon_0 > 0$. Here g and f are analytic functions in $|u-1| \le \varepsilon_0$ and satisfy $g(1) = 1$ and $f(1) = 1$. If

$$|\Xi_n(u)| = O\left(n^{1-\varepsilon}\right), \tag{28}$$

holds uniformly for $\{u \in \mathbb{C} : 1-\varepsilon_1 \le |u| \le 1+\varepsilon_1\} \setminus \{u \in \mathbb{C} : |u-1| \le \varepsilon_0\}$, where $0 < \varepsilon_1 < \varepsilon_0$, then we have, by standard application of the saddle-point method,

$$\mu_{n,k} = \frac{n}{\sqrt{2\pi\sigma^2 L_n}} e^{-\Delta^2/(2\sigma^2 L_n) + O\left(|\Delta|^3/L_n^2\right)} \left(1 + O\left(\frac{1+|\Delta|}{L_n}\right)\right), \tag{29}$$

uniformly for $|\Delta| \le L_n^{2/3}$, where $\Delta := k - f'(1)L_n$ and

$$\sigma = \sqrt{f'(1) + f''(1)}.$$

Note that to prove the estimate (29), we used (27) and (28) only when $u = e^{i\theta}$, $\theta \in \mathbb{R}$. However, the uniform estimates (27) and (28) in a complex neighborhood of unity also yield, by Cauchy's integral representation,

$$\mu_{n,k} = O\left(L_n^{-1/2} r^{-k} n^{f(r)} + r^{-k} n^{1-\varepsilon}\right) = O\left(L_n^{-1/2} r^{-k} n^{f(r)}\right), \tag{30}$$

uniformly for all $k = 0, \ldots, n$, where $r = 1 + o(1)$. Although this estimate becomes too crude for $|k - f'(1)L_n| \ge \varepsilon L_n$, it is sufficient for our purposes and very useful in bounding all error terms involved.

Estimates for $\nabla_{n,k}$. For higher central moments of the profile, we consider the difference

$$\nabla_{n,k}(\mathbf{j}) := \sum_{1\le \ell\le h} \mu_{j_\ell,k-1} - \mu_{n,k} \qquad (j_1+\cdots+j_h = n-\kappa), \tag{31}$$

where $\mathbf{j}=(j_1,\dots,j_h)$, which, by Cauchy's integral formula and (27), satisfies

$$\nabla_{n,k}(\mathbf{j}) = \frac{1}{2\pi i}\int_{\substack{|u|=r\\ |u-1|\le\varepsilon}} g(u)u^{-k-1}n^{f(u)}\varphi(u;\mathbf{j}/n)\left(1+O\left(\sum_{1\le\ell\le h}\frac{1}{j_\ell+1}\right)\right)\mathrm{d}u \\ + O\left(r^{-k}n^{1-\varepsilon}\right),$$

where

$$\varphi(u;\mathbf{j}/n) := \sum_{1\le\ell\le h} u\left(\frac{j_\ell}{n}\right)^{f(u)} - 1.$$

Since $\sum_{1\le\ell\le h} j_\ell = n+O(1)$, we deduce, by expanding $\varphi(u;x)$ at $u=1$, the two estimates

$$\nabla_{n,k}(\mathbf{j}) = \begin{cases} O\left(\left(|r-1|+L_n^{-1/2}\right)L_n^{-1/2}r^{-k}n^{f(r)}\right), & (r=1+o(1)),\\ O\left(|\Delta|L_n^{-3/2}n\right), \end{cases} \tag{32}$$

where the first estimate holds uniformly for all tuples $(j_1,\dots,j_h)$ and $1\le k<n$, and the second for all tuples $(j_1,\dots,j_h)$ and $\Delta = O(\sqrt{L_n})$. Note that if we take r to be the solution near unity of the equation $rf'(r) = k/L_n = f'(1)+\Delta/L_n$, then $r = 1+\Delta/(\sigma^2 L_n) + O(\Delta^2/L_n^2)$ and

$$r^{-k}n^{f(r)} = ne^{-\Delta^2/(2\sigma^2 L_n)+O(|\Delta|^3L_n^{-2})}, \tag{33}$$

uniformly for $\Delta = O(L_n^{2/3})$. This means that the first estimate in (32) is not tight when $\Delta = o(\sqrt{L_n})$.

Asymptotics of width and estimates needed for higher central moments. If we can prove that

$$\mathbb{E}\left\{(Y_{n,k}-\mu_{n,k})^m\right\} = O\left(|\Delta|^m L_n^{-m}\mu_{n,k}^m\right) \qquad (m\ge 0), \tag{34}$$

uniformly for $|\Delta| = o(L_n)$, then the width $W_n := \max_k Y_{n,k}$ satisfies the following estimates.

$$\begin{cases} \mathbb{E}\{W_n\} = \dfrac{n}{\sqrt{2\pi\sigma^2 L_n}}\left(1+O\left(L_n^{-1}\right)\right), \\ \mathbb{E}\left\{|W_n-\mathbb{E}\{W_n\}|^s\right\} = O\left(n^s L_n^{-3s/2}\right) \quad (s\ge 0), \\ \dfrac{W_n}{\mathbb{E}\{W_n\}} \longrightarrow 1 \quad \text{almost surely}, \\ \mathbb{P}\left(|k^*-f'(1)L_n|\ge T\right) = O\left(T^{-B}\right). \end{cases} \tag{35}$$

the last estimate holding for every $B>0$ and $T>T_0$, for some $T_0>1$.

Remarks for (35).

1. For the almost sure convergence in (35), we also need the additional property that *inserting a new node to the tree changes the width by at most a bounded quantity*, which is easily justified for almost all trees we are considering.

2. For most log trees we encounter, the O-estimate in the asymptotic approximation for the mean width in (35) is indeed tight; for simplicity, we content ourselves with the O-estimates.

3. From the preceding rough description, we see that the three main hard parts are (27), (28) and (34).

Instead of formulating a general theorem, which will be heavy at this stage, we consider in the following three major classes of log trees.

5 Random quadtrees and grid-trees

We start from quadtrees, and then indicate the estimates needed for the more general grid-trees proposed in Devroye (1998).

Construction of random quadtrees. Given a sequence of n points independently and uniformly chosen from $[0, 1]^d$, the random (point) quadtree associated with this random sample is constructed by placing the first point at the root, which splits the space into 2^d hyper-rectangles, each corresponding to one of the 2^d subtrees of the root. Points falling in each hyper-rectangle are directed to the corresponding subtree and are constructed recursively. For more information on quadtrees, see Flajolet et al. (1995), Chern et al. (2005) and the references therein.

The profile. By such a construction, the profile $Y_{n,k}$ satisfies (25) with $h = 2^d$ and

$$\begin{aligned}
\pi_{n,\mathbf{j}} &:= \mathbb{P}\left(I_{n,1} = j_1, \dots, I_{n,2^d} = j_{2^d}\right) \\
&= \binom{n-1}{j_1, \dots, j_{2^d}} \int_{[0,1]^d} \prod_{\substack{1\le \ell \le 2^d \\ \ell-1=(b_1,\dots,b_d)_2}} \left(\prod_{1\le i\le d} b_i(1-x_i) + (1-b_i)x_i \right)^{j_\ell} \mathrm{d}\mathbf{x},
\end{aligned}$$

where $(b_1, \dots, b_d)_2$ denotes the binary representation of $\ell-1$ (prefixed by zeros if $\lfloor \log_2(\ell-1) \rfloor < d-1$) and $\mathrm{d}\mathbf{x} = \mathrm{d}x_1 \cdots \mathrm{d}x_d$.

The underlying recurrence. From the expression for $\pi_{n,\mathbf{j}}$, it follows that all moments of $Y_{n,k}$ satisfy (26) with

$$\pi_{n,j} = \frac{1}{n} \sum_{j<j_1\le \cdots \le j_{d-1}\le n} \frac{1}{j_1 \cdots j_{d-1}} \qquad (0 \le j < n). \tag{36}$$

In particular, the expected profile $\mu_{n,k}$ satisfies the estimates (27) and (28) with $f(u) = 2u^{1/d} - 1$ and

$$g(u) := \frac{1}{\Gamma(2u^{1/d})^d(2u^{1/d}-1)} \prod_{1\le \ell<d} \frac{\Gamma(2u^{1/d}(1-e^{2\ell\pi i/d}))}{\Gamma(2-2u^{1/d}e^{2\ell\pi i/d})};$$

see Chern et al. (2005). The exact form of g is less important for our purpose; the analyticity of g for u near unity is however technically useful. Note that

$$f'(1) = \frac{2}{d}, \quad \sigma = \frac{\sqrt{2}}{d}.$$

Recurrence of $P_{n,k}^{(m)} := \mathbb{E}\{(Y_{n,k} - \mu_{n,k})^m\}$. Obviously, $P_{n,k}^{(0)} = 1$, $P_{n,k}^{(1)} = 0$ and $P_{n,k}^{(m)}$ satisfies the recurrence

$$P_{n,k}^{(m)} = 2^d \sum_{1\le j<n} \pi_{n,j} P_{j,k-1}^{(m)} + Q_{n,k}^{(m)} \qquad (m \ge 2),$$

where

$$Q_{n,k}^{(m)} := \sum_{(i_0,\dots,i_{2^d})\in\mathcal{I}_m} \binom{m}{i_0,\cdots,i_{2^d}} \sum_{j_1+\cdots+j_{2^d}=n-1} \pi_{n,\mathbf{j}} P_{j_1,k-1}^{(i_1)} \cdots P_{j_{2^d},k-1}^{(i_{2^d})} \nabla_{n,k}^{i_0}(\mathbf{j}). \tag{37}$$

Here $\nabla_{n,k}(\mathbf{j})$ is given in (31) with $h = 2^d$ and $\kappa = 1$ there and

$$\mathcal{I}_m := \{(i_0,\dots,i_{2^d}) \in \{0,\dots,m\}^{d+1} \,:\, 0 \le i_1,\dots,i_{2^d} < m\}.$$

Following the proof pattern for recursive trees, we prove, based on the estimates (32), the two bounds

$$P_{n,k}^{(m)} = \begin{cases} O\left(\left(|r-1|^m + L_n^{-m/2}\right) L_n^{-m/2} r^{-mk} n^{mf(r)}\right) & r = 1 + o(1) \\ O\left(|\Delta|^m L_n^{-3m/2} n^m\right), & \end{cases} \tag{38}$$

the first being uniform for $1 \le k < n$ and the second for $\Delta := k - f'(1)L_n = O(\sqrt{L_n})$. These two estimates imply (34) by (30) and (33).

An asymptotic transfer for the double-indexed recurrence. To justify the results in (35), it remains to prove the two estimates in (38). Note that an exact solution for (26) similar to (15) is still possible by the Euler-transform approach used in Flajolet et al. (1995), but the resulting expression is less manageable and the approach is less useful for other random log trees. Thus we use an inductive argument, which is easily amended for other varieties of trees.

Lemma 1. *Assume that $a_{n,k}$ satisfies (26) with $\pi_{n,j}$ given in (36) and $a_{n,0}, a_{1,k} = O(1)$. If*

$$|b_{n,k}| \le c|k - f'(1)L_n|^\lambda L_n^\beta \rho^{-k} n^\alpha,$$

for $n \ge 1$ and $1 \le k \le n$, where $\lambda \ge 0$, $\beta \in \mathbb{R}$, $c > 0$, and the two real numbers $\alpha, \rho > 0$ satisfy $\rho < ((\alpha+1)/2)^d$, then

$$|a_{n,k}| \le C_0|k - f'(1)L_n|^\lambda L_n^\beta \rho^{-k} n^\alpha, \tag{39}$$

for $n \ge 1$ and $1 \le k \le n$, where $C_0 > 0$ is chosen so large that $C_0 \ge c/\left(1 - \frac{2^d\rho}{(\alpha+1)^d} - \varepsilon\right)$.

Proof. We apply induction on k and n. The boundary conditions are easily checked by taking C_0 sufficiently large. We may assume that $|k - f'(1)L_n| \to \infty$, for otherwise we need only modify the value of c. By induction hypothesis, we have (see Chern et al., 2005)

$$\begin{aligned}
|a_{n,k}| &\le c|k - f'(1)L_n|^\lambda L_n^\beta \rho^{-k} n^\alpha \\
&\quad + 2^d C_0 \rho^{1-k} n^{-1} \sum_{1\le j<j_1\le\cdots\le j_{d-1}\le n} \frac{|k-1-f'(1)L_j|^\lambda L_j^\beta j^\alpha}{j_1\cdots j_{d-1}} \\
&\sim c|k - f'(1)L_n|^\lambda L_n^\beta \rho^{-k} n^\alpha \\
&\quad + \frac{2^d C_0}{(d-1)!}\rho^{1-k} n^{-1} \sum_{1\le j<n} |k - f'(1)L_j|^\lambda L_j^\beta j^\alpha \left(\log\frac{n}{j}\right)^{d-1} \\
&= cn^\alpha L_n^\beta \rho^{-k} + \frac{2^d}{(\alpha+1)^d} C_0(1+o(1))|k - f'(1)L_n|^\lambda L_n^\beta \rho^{1-k} n^\alpha;
\end{aligned}$$

thus (39) follows by properly tuning C_0 (since $\rho < ((\alpha+1)/2)^d$). □

Asymptotics of $P_{n,k}^{(m)}$. We prove first by induction the first bound in (38). Assume $m \ge 2$.

Consider $Q_{n,k}^{(m)}$. We split the inner sum in (37) similar to the $Q_{n,k}^{(m)}$ of recursive trees. If $j_1, \dots, j_{2^d} \ge n/L_n^m$, then $L_{j_\ell} \sim L_n$ for $\ell = 1, \dots, 2^d$, and we have

$$\begin{aligned}
&\sum_{(i_0,\dots,i_{2^d})\in\mathcal{I}_m} \binom{m}{i_0,\dots,i_{2^d}} \sum_{n/L_n^m \le j_1,\dots,j_{2^d}<n} \pi_{n,\mathbf{j}} P_{j_1,k-1}^{(i_1)} \cdots P_{j_{2^d},k-1}^{(i_{2^d})} \nabla_{n,k}^{i_0}(\mathbf{j}) \\
&\quad = O\left(\left(|r-1|^m + L_n^{-m/2}\right) L_n^{-m/2} r^{-mk} n^{mf(r)}\right). \qquad (40)
\end{aligned}$$

We now assume that one of the j_ℓ's, say j_1, is less than n/L_n^m. We may furthermore assume that the corresponding index i_1 of j_1 is nonzero; for otherwise, if all $i_\ell = 0$ for those j_ℓ's with $j_\ell \le n/L_n^m$, then the bound on the right-hand side of (40) obviously holds since all other j_ℓ's satisfy $L_{j_\ell} \sim L_n$. Terms in $Q_{n,k}^{(m)}$ with $i_1 \ge 1$ and $j_1 \le n/L_n^m$ are bounded above by

$$O\left(r^{-mk} \sum_{(i_0,\dots,i_{2^d})\in\mathcal{I}_m} n^{(m-i_1)f(r)} \sum_{j_1\le n/L_n^m} \pi_{n,j_1} j_1^{i_1 f(r)}\right) = O\left(L_n^{-m} r^{-mk} n^{mf(r)}\right).$$

This proves that

$$Q_{n,k}^{(m)} = O\left(\left(|r-1|^m + L_n^{-m/2}\right) L_n^{-m/2} r^{-mk} n^{mf(r)}\right).$$

Thus the first estimate in (38) holds by applying the O-transfer of Lemma 1.

The proof of the second estimate in (38) follows the same inductive argument, details being omitted here.

Consequently, the width W_n of random quadtrees satisfies all approximations in (35); in particular, the expected width satisfies

$$\mathbb{E}\{W_n\} = \frac{dn}{2\sqrt{\pi L_n}}\left(1 + O\left(L_n^{-1}\right)\right).$$

All our results are new except when $d = 1$ for which quadtrees reduce to binary search trees and the almost sure convergence in (35) was derived in Chauvin et al. (2001), and the expected width in Drmota and Hwang (2005b) (with a weaker error term).

Random grid-trees. Grid-trees were first proposed by Devroye (1998) and represent one of the extensions of quadtrees. Instead of placing the first element in the given sequence at the root (as in quadtrees), we fix an integer $m \geq 2$ and place the first $m-1$ elements at the root, which then split the space into m^d hyper-rectangles (called grids). The remaining construction is similar to that for quadtrees.

In this case, we have $h = m^d$ and ($j_0 := j$, $j_d := n-m+1$)

$$\pi_{n,j} = \sum_{j \leq j_1 \leq \cdots \leq j_{d-1} \leq n-m+1} \prod_{1 \leq \ell \leq d} \frac{\binom{j_\ell - j_{\ell-1}+m-2}{m-2}}{\binom{j_\ell+m-1}{m-1}},$$

and (27) and (28) hold by applying the approach proposed in Chern et al. (2005), where $f(u)$ satisfies

$$((f(u)+1)\cdots(f(u)+m-1))^d = m!^d u \qquad (m \geq 2; d \geq 1),$$

with $f(1) = 1$. An O-transfer similar to that given in Lemma 1 can also be derived by noting that

$$\begin{aligned}
&\sum_{1 \leq j < n} \pi_{n,j} |k - 1 - f'(1) L_j|^\lambda L_j^\beta j^\alpha \\
&= \sum_{1 \leq j \leq j_1 \leq \cdots \leq j_{d-1} \leq n-m+1} |k-1-f'(1)L_j|^\lambda L_j^\beta j^\alpha \prod_{1 \leq \ell \leq d} \frac{\binom{j_\ell - j_{\ell-1}+m-2}{m-2}}{\binom{j_\ell+m-1}{m-1}} \\
&\sim \frac{(m-1)^d}{n} \sum_{1 \leq j_{d-1} \leq n} \frac{1}{j_{d-1}} \left(1 - \frac{j_{d-1}}{n}\right)^{m-2} \sum_{1 \leq j_{d-2} \leq j_{d-1}} \cdots \\
&\qquad \cdots \times \sum_{1 \leq j_1 \leq j_2} \frac{1}{j_1} \left(1 - \frac{j_1}{j_2}\right)^{m-2} \sum_{1 \leq j \leq j_1} |k - f'(1)L_j|^\lambda L_j^\beta j^\alpha \left(1 - \frac{j}{j_1}\right)^{m-2} \\
&\sim (m-1)^d \left(\frac{\Gamma(m-1)\Gamma(\alpha+1)}{\Gamma(m+\alpha)}\right)^d |k - f'(1)L_n|^\lambda L_n^\beta n^\alpha;
\end{aligned}$$

so that the same type of asymptotic transfer there holds when $\alpha, \rho > 0$ satisfy the inequality

$$\rho < \left(\frac{(\alpha+1)\cdots(\alpha+m-1)}{m!}\right)^d.$$

Then the approximations (35) hold for the width and we have

$$f'(1) = \frac{1}{d(H_m - 1)}, \quad \sigma = \sqrt{\frac{H_m^{(2)} - 1}{d^2(H_m-1)^3}},$$

where $H_m := \sum_{1 \leq j \leq m} 1/j$ and $H_m^{(2)} := \sum_{1 \leq j \leq m} 1/j^2$. Note that $d = 1$ corresponds to m-ary search trees, and $m = 2$ to quadtrees. No martingale structure is known for grid-trees for general (m, d). Our results are new.

6 Generalized m-ary search trees

Binary search trees, which are special cases of quadtrees and m-ary search trees, have yet another extension; see Hennequin (1991), Chern and Hwang (2001a). Instead of placing the first $m-1$ elements in the given sequence of numbers at the root (as in m-ary search trees), we choose a random sample of $m(t+1)-1$ elements, where $m \geq 2$ and $t \geq 0$, and sort it in increasing order. Then use the $(t+1)$-st, the $2(t+1)$-st, …, and the $(m-1)(t+1)$-st smallest elements in the sample to partition the original sample into m groups, corresponding to the m subtrees of the root. Elements falling in each subtree are constructed recursively in the same way and the process stops as long as the subtree size is less than $m(t+1)-1$, which can then be arranged arbitrarily since asymptotically this will have limited effect.

In this case, the profile $Y_{n,k}$ satisfies (25) with $h = m$ and

$$\mathbb{P}(I_{n,1} = j_1, \cdots, I_{n,m} = j_m) = \frac{\binom{j_1}{t} \cdots \binom{j_m}{t}}{\binom{n}{m(t+1)-1}}.$$

Furthermore, (27) and (28) hold with $f(u)$ satisfying the equation

$$(f(u)+t+1)\cdots(f(u)+m(t+1)-1) = \frac{(m(t+1))!}{(t+1)!}\,u,$$

with $f(1) = 1$, where $m \geq 2$ and $t \geq 0$; see Chern and Hwang (2001a, 2001b) for the asymptotic tools needed (based on differential equations). Straightforward computation gives

$$f'(1) = \frac{1}{H_{m(t+1)} - H_{t+1}}, \quad \sigma = \sqrt{\frac{H^{(2)}_{m(t+1)} - H^{(2)}_{t+1}}{(H_{m(t+1)} - H_{t+1})^3}}.$$

The estimate (34) can be checked by an inductive argument similar to quadtrees using the expression

$$\pi_{n,j} = \frac{\binom{j}{t}\binom{n-1-j}{(m-1)(t+1)-1}}{\binom{n}{m(t+1)-1}}.$$

In particular, we can derive an O-transfer similar to Lemma 1 with the two numbers α, ρ there satisfying

$$\rho < \frac{(\alpha+t+1)\cdots(\alpha+m(t+1)-1)}{(t+2)\cdots(m(t+1))}.$$

Note that m-ary search trees correspond to $t = 0$, and $m = 2$ reduces to the so-called fringe-balanced or median-of-$(2t+1)$ binary search trees; see Devroye (1998).

7 Random increasing trees

Increasing trees are rooted, labeled trees with labels along any path down from any node forming an increasing sequence; see Bergeron et al. (1992). The exponential generating functions $\tau(z) := \sum_n \tau_{n\geq 1} z^n/n!$ for the number τ_n of increasing trees often has the form

$$\tau'(z) = \phi(\tau(z)), \tag{41}$$

with $\tau(0) = 0$ and $\tau(1) = 1$, for some function $\phi(w)$ with $\phi(0) = 1$ and nonnegative coefficients. In this case, there are three representative varieties of increasing trees: (i) recursive trees with $\phi(w) = e^w$; (ii) binary increasing trees with $\phi(w) = (1 + w)^2$, and (iii) plane-oriented recursive trees (*PORTs*) with $\phi(w) = 1/(1 - w)$.

We already studied the width of random recursive trees and random binary increasing trees (identically distributed as random binary search trees). We consider first PORTs and then mention other varieties of increasing trees (in some generality).

Random PORTs. PORTs are labeled, ordered (or plane) trees with the property that labels along any path down from the root are increasing; see Bergeron et al. (1992), Prodinger (1995) for more details.

The recurrence for the profile $Y_{n,k}$ is similar to (8) but with a very different underlying distribution (see Hwang, 2005)

$$Y_{n,k} \stackrel{d}{=} Y_{I_n,k-1} + Y^*_{n-I_n,k} \qquad (n \geq 2; k \geq 1),$$

where the $Y^*_{n,k}$'s are independent copies of $Y_{n,k}$ and

$$\pi_{n,j} = \mathbb{P}(I_n = j) = \frac{2\binom{2j-2}{j-1}\binom{2n-2j-2}{n-j-1}}{j\binom{2n-2}{n-1}} \qquad (1 \leq j < n).$$

We have

$$\Xi_n(u) = \frac{1}{(1+u)}\left(\frac{2\sqrt{\pi}}{\Gamma(u/2)} n^{(u+1)/2} + 1\right)\left(1 + O\left(n^{-\varepsilon}\right)\right),$$

uniformly for $|u| \leq C$, for any $C > 0$; see Bergeron et al. (1992), Hwang (2005), Prodinger (1996). Thus $f(u) = (u + 1)/2$, so that $\sigma = 1/\sqrt{2}$. All approximations in (35) hold for the width. Note that although the recurrence satisfied by $Y_{n,k}$ is not of the form (25), the technicalities are similar to those for recursive trees; see Hwang (2005) for details.

The widths and profiles of random increasing trees for which $1/\phi(w)$ equals a polynomial also exhibit similar behaviors.

Polynomial varieties. We now show that the same results (35) (except the almost sure convergence) also hold for polynomial varieties of increasing trees; see Bergeron et al. (1992). Briefly, these are increasing trees whose exponential generating function $\tau(z) := \sum_{n\geq 1} \tau_n z^n/n!$ satisfies (41) with

$$\phi(w) := \sum_{0\leq j\leq d} \phi_j w^j \qquad (d \geq 2),$$

where $\phi_j \geq 0$ for $0 \leq j \leq d$ and $\phi_0, \phi_d > 0$. In this case, it is known that

$$\frac{\tau_n}{n!} = \frac{p}{\Gamma(\frac{1}{d-1})}\left((d-1)\phi_d R\right)^{-1/(d-1)} R^{-n} n^{-(d-2)/(d-2)}\left(1 + O\left(n^{-2/(d-1)}\right)\right), \qquad (42)$$

where p denotes the period of $\phi(v)$, $R := \int_0^\infty \mathrm{d}v/\phi(v)$ and

$$\sum_{n,k} \mathbb{E}\{Y_{n,k}\} u^k \frac{z^n}{n!} = \left(\tau'(z)\right)^{-u} \int_0^z \left(\tau'(v)\right)^{1-u} \mathrm{d}v;$$

see Bergeron et al. (1992). From these relations, we deduce the two estimates (27) and (28) with

$$f(u) = \frac{d}{d-1}u - \frac{1}{d-1},$$
$$g(u) = \phi_d^{(1-w)/(d-1)} (R(d-1))^{(1-dw)/(d-1)} \frac{\Gamma(1/(d-1))}{\Gamma(dw/(d-1))} \int_0^\infty \phi(v)^{-w} \mathrm{d}v.$$

Furthermore, the higher moments of $Y_{n,k}$ (centered or not) satisfy the recurrence

$$a_{n,k} = b_{n,k} + \sum_{1 \le j < n} \varpi_{n,j} a_{j,k-1},$$

where

$$\varpi_{n,j} := \frac{(n-1)!\tau_j}{\tau_n j!}[z^{n-1-j}]\phi'(\tau(z)).$$

By (42), we then derive an O-transfer for $a_{n,k}$ similar to Lemma 1 with α and ρ there satisfying

$$\rho < \frac{d-1}{dp^{d-1}}\left(\alpha + \frac{1}{d-1}\right);$$

from this the estimates (34) are then justified.

Random mobile trees. These are increasing trees whose enumerating generating function satisfies (41) with $\phi(w) = 1 - \log(1-w)$; see Bergeron et al. (1992). This example is less natural but very interesting because nodes are distributed in a rather different way. First, the expected profile is given by

$$\Xi_n(u) = \sum_k \mu_{n,k} u^k = \frac{n!}{\tau_n}[z^n]\tau'(z)^u \int_0^z \tau'(v)^{1-u} \mathrm{d}v.$$

Here the number τ_n of such trees satisfies

$$\frac{\tau_n}{n!} = R^{1-n} n^{-2} \left(1 + O\left(L_n^{-1}\right)\right),$$

where $R = \int_0^\infty (1+v)^{-1} e^{-v} \mathrm{d}v$. By singularity analysis (see Flajolet and Odlyzko, 1990), we deduce that

$$\Xi_n(u) = g(u) n L_n^{u-1} \left(1 + O\left(\frac{\log L_n}{L_n}\right)\right),$$

where the O-term holds uniformly for bounded complex and

$$g(u) = R^{-1} u \int_0^\infty e^{-v}(1+v)^{-u} \mathrm{d}v.$$

Note that this is not of the form (27) and $g(1) = 1$. Thus such mobile trees are very "bushy" at each level (the root already having about n/L_n nodes) and we have

$$\max_k \mu_{n,k} \sim \frac{n}{\sqrt{2\pi \log L_n}},$$

the mode being reached at $k \sim \log L_n$. The same methods of proof we used for recursive trees can be extended to show that

$$\mathbb{E}\{W_n\} \sim \frac{n}{\sqrt{2\pi \log L_n}},$$

a very different behavior from all types of random trees we have discussed.

References

[1] D. Aldous, The continuum random tree. II. An overview. In *Stochastic Analysis (Durham, 1990)*, 23–70, LMS Lecture Note Series, 167, Cambridge University Press, Cambridge, 1991.

[2] F. Bergeron, P. Flajolet, and B. Salvy (1992). Varieties of increasing trees. In *CAAP'92* (Rennes, 1992), pp. 24–48, *Lecture Notes in Computer Science*, 581, Springer, Berlin, 1992.

[3] P. Chassaing and J.-F. Marckert (2001). Parking functions, empirical processes, and the width of rooted labeled trees. *Electronic Journal of Combinatorics* **8**, no. 1, Research Paper 14, 19 pp. (electronic).

[4] B. Chauvin, M. Drmota and J. Jabbour-Hattab (2001). The profile of binary search trees. *Annals of Applied Probability* **11** 1042–1062.

[5] H.-H. Chern, M. Fuchs and H.-K. Hwang (2005). Phase changes in random point quadtrees. Manuscript submitted for publication; available at algo.stat.sinica.edu.tw.

[6] H.-H. Chern and H.-K. Hwang (2001a). Phase changes in random m-ary search trees and generalized quicksort. *Random Structures and Algorithms* **19** 316–358.

[7] H.-H. Chern and H.-K. Hwang (2001b). Transitional behaviors of the average cost of quicksort with median-of-(2t+1). *Algorithmica* **29** 44–69.

[8] H.-H. Chern, H.-K. Hwang and T.-H. Tsai (2003). An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms. *Journal of Algorithms* **44** 177–225.

[9] L. Devroye (1998). Universal limit laws for depths in random trees. *SIAM Journal on Computing* **28** 409–432.

[10] M. Drmota and B. Gittenberger (2004). The width of Galton-Watson trees conditioned by the size. *Discrete Mathematics and Theoretical Computer Science* **6** 387–400.

[11] M. Drmota and H.-K. Hwang (2005a). Bimodality and phase transitions in the profile variance of random binary search trees. *SIAM Journal on Discrete Mathematics* **19** 19–45.

[12] M. Drmota and H.-K. Hwang (2005b). Profiles of random trees: correlation and width of random recursive trees and binary search trees. *Advances in Applied Probability* **37** 321–341.

[13] P. Erdős (1953). On a conjecture of Hammersley. *Journal of the London Mathematical Society* **28** 232–236.

[14] P. Flajolet, G. Labelle, L. Laforest, and B. Salvy (1995). Hypergeometrics and the cost structure of quadtrees. *Random Structures and Algorithms* **7** 117–114.

[15] P. Flajolet and G. Louchard (2001). Analytic variations on the Airy distribution. *Algorithmica* **31** 361–377.

[16] P. Flajolet and A. M. Odlyzko (1990). Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics* **3** 216–240.

[17] M. Fuchs, H.-K. Hwang and R. Neininger (2005). Profiles of random trees: Limit theorems for random recursive trees and binary search trees. Manuscript submitted for publication; available at algo.stat.sinica.edu.tw.

[18] J. M. Hammersley (1951). The sum of products of the natural numbers. *Proceedings of the London Mathematical Society* **1** 435–452.

[19] P. Hennequin (1991). *Analyse en moyenne d'algorithme, tri rapide et arbres de recherche*. Ph.D. Thesis, LIX, École polytechnique.

[20] H.-K. Hwang (1994). *Théorèmes limites pour les structures combinatoires et les fonctions arithmétiques*. Ph. D. Thesis, LIX, École polytechnique.

[21] H.-K. Hwang (1995). Asymptotic expansions for the Stirling numbers of the first kind. *Journal of Combinatorial Theory, Series A* **71** 343–351.

[22] H.-K. Hwang (2005). Profiles of random trees: plane-oriented recursive trees. Manuscript submitted for publication; available at algo.stat.sinica.edu.tw.

[23] M. Loève (1977). *Probability theory, I*. Fourth edition, Springer-Verlag, NY.

[24] H. Prodinger (1996). Depth and path length of heap ordered trees. *International Journal of Foundations of Computer Science* **7** 293–299.