行政院國家科學委員會專題研究計畫　成果報告

隨機樹狀之機率性質

計畫類別：個別型計畫
計畫編號：NSC93-2119-M-009-003-
執行期間：93 年 10 月 01 日至 94 年 10 月 31 日
執行單位：國立交通大學應用數學系(所)

計畫主持人：符麥克

報告類型：精簡報告
報告附件：出席國際會議研究心得報告及發表論文
處理方式：本計畫可公開查詢

中　華　民　國 94 年 11 月 28 日

成果報告

符麥克

計畫名稱：隨機樹狀之機率性質

摘要：

在這個計畫中，我們研究隨機樹的機率性質。本質上，我們專注在兩個種類的樹：遞迴樹和四分樹。我們設計一個一般性的方法並得到遞迴樹的輪廓(profile)的極限法則。此外，我們給出一個可用於導出四分樹的特徵參數如葉子的數目、總路徑長度等等隨機性質的一般性構造。這個新的方法可讓我們重新得到大多數以前的結果和很多新的結果。這兩種方法的共同之處在於他們依據 method of moments 和漸進轉換定理。而且，他們都有一般性並且可預期地有著很多的應用。

Project: Probabilistic properties of random trees

by

Michael Fuchs

Abstract

In this project, we investigated probabilistic properties of random trees. We essentially focused on two types of trees: recursive trees and quadtrees. We devised a general method to derive limit laws for the profile of recursive trees. Furthermore, we gave a general framework to derive stochastic properties for a large class of characteristic parameters such as the number of leaves, the total path lengths, etc. of quadtrees. This new approach allowed us to re-derive most of the previous results and to add many new ones. What both methods have in common is that they rest on the method of moments and on asymptotic transfer theorems. Moreover, they are both of some generality and are expected to have many more applications.

1 General

This is the final report on the National Science Council project entitled "Probabilistic properties of random trees" with number 93-2119-M-009-003 and term from November 1st, 2004 to October 31st, 2005.

Before presenting our results in more details, we give an overview of the main outcomes of the project.

- Parts of the paper [4] were written within this project. The paper was submitted and is accepted for publication in one of the forthcoming issues of *Algorithmica*.
- The paper [1] was written within this project. It was submitted and is currently under review.
- The results of the second paper were presented at the 12th International Conference on Random Structures and Algorithms in Poznan, Poland (a report on the conference was already handled in at an earlier stage).

2 Results

The aim of the project was to study the stochastic behavior of characteristic parameters of random trees. In particular, we were interested in the so called profile which roughly speaking is the shape of the tree. Subsequently, we briefly describe the main results of the to papers [4] and [1] which were written within this project.

1. Profiles of random trees: Limit theorems for random recursive trees and binary search trees. A manuscript of this paper existed already at the time of the proposal of the project. Actually, the results were already discussed when applying for the project (see the project proposal). It was one of the original goals of the project to extend the results to m-ary search trees and median-of-$2t+1$ search trees.

As already discussed in the project proposal the main complications of such an extension are arising from the more technical nature of the latter two families of random search trees. When trying to extend the original method, we realized that it can be largely simplified. [4] is an improved version of the earlier manuscript using a more simplified method of proof.

In order to provide some details of where the simplifications occur, denote by $X_{n,k}$ the number of external nodes at level k of a random binary search tree on n nodes. An exemplary result of [4] reads then as follows.

Theorem 1. *Assume that $k = \alpha \log n + o(\log n)$, where $\alpha \in [1, 2]$. Then,*

$$\frac{X_{n,k}}{\mathbb{E}X_{n,k}} \xrightarrow{d} X_\alpha,$$

with convergence of all moments.

For the proof, we used the method of moments. Therefore, we had to study the transfer behavior of the following recurrence

$$a_{n,k} = \frac{2}{n} \sum_{j=0}^{n-1} a_{j,k-1} + b_{n,k},$$

which arises when studying the moments and central moments of $X_{n,k}$. Here, $b_{n,k}$ is a given sequence. The transfer behavior of such and similar recurrences was already studied in previous works. The new feature of the present study is the dependence on two indices n, k which makes the problem more involved.

From such transfer theorems the above result can then be obtained by proving the following asymptotic expansions for all moments

$$\mathbb{E}X_{n,k}^m \sim \nu_m(\alpha) \left(\mathbb{E}X_{n,k}\right)^m,$$

where $k = \alpha \log n + o(\log n)$ and $\nu_m(\alpha) = \mathbb{E}X_\alpha^m$. The proof of the latter proceeds in two steps.

1. Obtain an upper bound for $\mathbb{E}X_{n,k}^m$ uniformly valid for all n, k.
2. For $k = \alpha \log n + o(\log n)$ refine the analysis of the previous step.

In our previous manuscript of [4], the first step above was technically involved. We succeeded in finding a simpler approach, thereby greatly simplifying the original method.

This new approach is expected to work as well for m-ary search trees and median-of-$2t + 1$ search trees. In particular, it will make the expected technical difficulties for these more complicated random search tree structures easier to handle. This is work still in progress and might be the topic of a forthcoming project.

2. Phase changes in random point quadtrees. Apart from studying the profile of m-ary search trees and median-of-$2t + 1$ search trees, also the profile of quadtrees (yet another extension of binary search trees) is of great interest. The paper [1] is expected to provide the technical machinery needed for such a study. However, the results of the paper are also of great interest on their own.

A future detailed study of the profile was only one driving motivation for the techniques developed in [1]. Another source of inspiration was a recent paper of Dean and Majumdar [2] were they observed a phase transition in the random continuous fragmentation problem: the limit law changes from normal which they could prove rigorously to non-normal which they just concluded from experiments. Since their model is closely related to random quadtrees a similar behavior is expected to hold for the number of leaves X_n of a random point quadtree of dimension d.

The number of leaves for $d = 1$ (this corresponds to the binary search tree) was well-studied in literature and the situation is well-understood. However, for general d, only precise asymptotic expansions for the mean value were derived by Flajolet et al. [3] previous to our work. More precisely, they proved that

$$\mathbb{E}X_n = n - \sum_{2\le k\le n} \binom{n}{k}(-1)^k[k]! \sum_{2\le j\le k} \frac{1}{[j]!},$$

where $[k]! := \prod_{3\le j\le k}\left(1-2^d/j^d\right)$ for $k \ge 3$ and $[2]! := 1$. From this explicit expression, asymptotic expansions can be quickly derived by using Rice's integral method.

In [1] we devised a general method to obtain asymptotic expansions for all central moments. Such results then quickly entail the following theorem confirming the results obtained heuristically in [2].

Theorem 2. *(i) If $1 \le d \le 8$ then*

$$\frac{X_n - \mu_d n}{\sigma_d\sqrt{n}} \to N(0,1),$$

where $N(0,1)$ denotes the standard normal law and μ_d, σ_d are suitable constants.

(ii) If $d > 8$ then $(X_n - \mathbb{E}(X_n))/\sqrt{\mathbb{V}(X_n)}$ does not converge to a fixed limit law.

Moreover, we refined our method in order to get the following result which completely clarifies the second phase change.

Theorem 3. *Let*

$$\bar{\alpha} := \begin{cases} 1/3, & \text{if } 1 \le d \le 7; \\ \sqrt{2}-1, & \text{if } d = 8. \end{cases}$$

Then,

$$\mathbb{P}\left(X_n = \lfloor \mathbb{E}(X_n) + x\sqrt{\mathbb{V}(X_n)}\rfloor\right) = \frac{e^{-x^2/2}}{\sqrt{2\pi\mathbb{V}(X_n)}}\left(1 + \mathcal{O}\left(\left(1+|x|^3\right)n^{-3(1/2-\bar{\alpha})}\right)\right)$$

and

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} < x\right) - \Phi(x)\right| = \mathcal{O}\left(n^{-3(1/2-\bar{\alpha})}\right),$$

where the above rate is optimal. Here, $\Phi(x)$ denotes the distribution function of the standard normal distribution.

The proofs of these two results again rest on the method of moments and its refinement. As already explained above, the main step is to study the asymptotic transfer behavior of the underlying recurrence for the (centralized) moments which here has the form

$$a_n = 2^d \sum_{j=0}^{n-1} \pi_{n,j} a_j + b_n,$$

where

$$\pi_{n,j} = \binom{n-1}{j} \int_{[0,1]^d} (x_1\cdots x_d)^j (1 - x_1\cdots x_d)^{n-1-j} \mathrm{d}\mathbf{x}.$$

By introducing generating functions, the above recurrence can be translated into a differential equation. Where in most previous studies the so obtained differential equations was of Cauchy-Euler type, the present situation is complicated by the fact that the differential equation is not of Cauchy-Euler type anymore. However, the differential equation can be interpreted as a perturbed differential equation of Cauchy-Euler type. In [1], we developed the machinery to obtain asymptotic transfer theorems for such differential equations which then also yield corresponding asymptotic transfer theorems for the above recurrence. We omit stating detailed results here and instead direct the interested reader to [1].

Our method is general enough to be applicable to a wide range of characteristic parameters of quadtrees, thereby re-deriving most of the previous results and adding many new ones. Just to give some more examples, in addition to the number of leaves discussed above, the methods yields a variety of results such as precise asymptotic expansions of moments, first and second phase change, etc. as well for

- Paging;
- Node sorts;
- Total path lengths;
- Expected profile;

and many more.

Moreover the method can be applied to derive similar results for Devroye's random grid trees which constitute a common extension of both quadtrees and m-ary search trees.

For more details and more results, the reader may directly consult the paper [1].

As already mentioned in the introduction, the techniques we developed for quadtrees are expected to be applicable as well to derive finer results (in the flavor of [4]) for the profile. Of course, due to the more complicated nature of the profile, some further technical complications are expected. This is work in progress and might be the topic of an another project.

3 Summary

We shortly summarize the results of this project and indicate some future directions of research.

- A previous version of [4] was largely improved by greatly simplifying the method of proof. The new manuscript was then submitted and is about to be published.
- We gave a general framework to study probabilistic properties of a huge class of characteristic parameters of quadtrees and grid trees, thereby re-deriving most of the previous results and adding many new ones.
- The two previous papers are expected to lay out the tools needed for a detailed study of the profile of $\log$-trees such as m-ary search trees, median-of-$2t+1$ search trees, quadtrees, etc. This might be the topic of a forthcoming project.
- The method devised in [4] is expected to have many more applications such as in the study of the number of subtrees of a given size in random search trees, the number of nodes of fixed outdegree, etc. This might be the topic of yet another forthcoming project.

References

[1] H.-H. Chern, M. Fuchs, H.-K. Hwang (2005). Phase changes in random point quadtrees, submitted.

[2] D. S. Dean and S. N. Majumdar (2002). Phase transition in a random fragmentation problem with applications to computer science, *Journal of Physics. A. Mathematical and General*, **35**, L501-507.

[3] P. Flajolet, G. Labelle, L. Laforest, B. Salvy (1995). Hypergeometrics and the cost structure of quadtrees, *Random Structures and Algorithms*, **7**, 117-144.

[4] M. Fuchs, H.-K. Hwang, R. Neininger (2005). Profiles of random trees: Limit theorems for random recursive trees and binary search trees, *Algorithmica*, accepted for publication.

Phase changes in random point quadtrees

Hua-Huai Chern[a]
Department of Computer Science
National Taiwan Ocean University
Keelung 202
Taiwan

Michael Fuchs[b]
Department of Applied Mathematics
National Chiao Tung University
Hsinchu 300
Taiwan

Hsien-Kuei Hwang[c]
Institute of Statistical Science
Academia Sinica
Taiwan

December 31, 2004

Dedicated to the memory of Ching-Zong Wei (1949–2004)

Abstract

We show that a wide class of linear cost measures (such as the number of leaves) in random d-dimensional point quadtrees undergo a change in limit laws: if the dimension $d = 1, \dots, 8$, then the limit law is normal; if $d \geq 9$ then there is no convergence to a fixed limit law. Stronger approximation results such as convergence rates and local limit theorems are also derived for the number of leaves, additional phase changes being unveiled. Our approach is new and very general, and also applicable to other classes of search trees. A brief discussion of Devroye's grid-trees (covering m-ary search trees and quadtrees as special cases) is given. We also propose an efficient numeric procedure for computing the constants involved to high precision.

1 Introduction

Phase transitions in random combinatorial objects issuing from computer algorithms have received much recent attention by computer scientists, probabilists, and statistical physists, especially for NP-complete problems. We address in this paper the change of the limit laws from normal to non-convergence of some cost measures in random point quadtrees when the dimension varies. The phase change phenomena[1], as well as the asymptotic tools we develop (based mostly on linear operators), are of some generality. We will discuss the corresponding phase changes in Devroye's random grid-trees (see [12]) for which a complete description of the phase changes will be given.

[a]Partially supported by National Science Council of ROC under the Grant NSC-93-2115-M-019-001.
[b]Partially supported by National Science Council of ROC under the Grant NSC-93-2119-M-009-003.
[c]Partially supported by a Research Award of the Alexander von Humboldt Foundation.

[1]We use mostly "phase change" instead of "phase transition" because the dimension in our problem takes only positive integers.

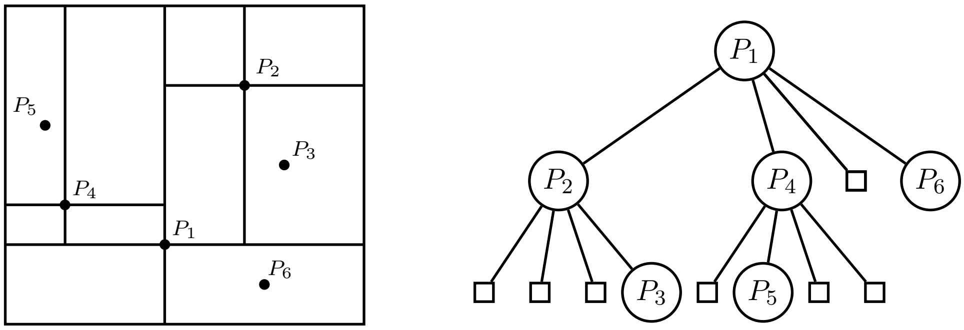

Figure 1: *A configuration of* 6 *points in the unit square and the corresponding quadtree.*

Point quadtrees. Point quadtrees, first introduced by Finkel and Bentley [16], are useful spatial and indexing data structures in computational geometry and for low-dimensional points in diverse applications in practice; see de Berg et al. [9], Samet [43, 44] for more information. In this paper, *we will say quadtrees instead of point quadtrees* for simplicity.

Given a sequence of points in $\mathbb{R}^d$, the quadtree associated with this point sequence is constructed as follows. The first point is placed at the root and then splits the underlying space into 2^d smaller regions (or quadrants), each corresponding to one of the 2^d subtrees of the root. The remaining points are directed to the quadrants (or the corresponding subtrees), and the subtrees are then constructed recursively by the same procedure. See Figure 1 for a plot of $d = 2$. When $d = 1$, quadtrees are simply binary search trees. Thus quadtrees can be viewed as one of the many different extensions of binary search trees; see [7, 12, 37].

Random quadtrees. To study the typical shapes or cost measures of quadtrees, we assume that the given points are uniformly and independently chosen from $[0, 1]^d$, where $d \geq 1$, and then construct the quadtree associated with the random sequence; the resulting quadtree is called a *random quadtree*.

Several shape parameters and cost measures in random quadtrees have been studied, reflecting in different levels certain typical complexity of algorithms on quadtrees.

- Depth (distance of a randomly chosen node to the root): [12, 13, 17, 19, 20];
- Total path length (sum of distances of all nodes to the root): [17, 19, 40];
- Cost of partial-match queries: [4, 17, 38, 41];
- Node types: [19, 26, 34, 35, 36];
- Height (distance of the longest path to the root): [10, 12].

In particular, the asymptotic normality of the depth was first proved in Flajolet and Lafforgue [20] (see also [12]), and the non-normal limit law for the total path length in Neininger and Rüschendorf [40].

The number of leaves. For concreteness and simplicity, we present the phase change phenomena through the number of leaves, denoted by $X_n = X_{n,d}$, in random quadtrees of n points. The extension to more general cost measures will be discussed later.

When $d = 1$, it is known that X_n (the number of leaves in random binary search trees of n nodes) is asymptotically normally distributed with mean and variance asymptotic to $n/3$ and $2n/45$, respectively; see [11, 18]. A local limit theorem is also given in [18].

For $d \geq 2$, Flajolet et al. (see [19]) first derived the closed-form expression for the expected value of X_n

$$\mathbb{E}(X_n) = n - \sum_{2\leq k\leq n} \binom{n}{k} (-1)^k [k]! \sum_{2\leq j\leq k} \frac{1}{[j]!} \qquad (n \geq 1), \tag{1}$$

where $[k]! := \prod_{3\leq j\leq k}(1 - 2^d/j^d)$ for $k \geq 3$ and $[2]! := 1$, and then showed that

$$\mathbb{E}(X_n) \sim \mu_d n,$$

where

$$\mu_d := 1 - \frac{2}{d} \prod_{\ell\geq 3} \frac{1}{1 - \left(\frac{2}{\ell}\right)^d} + 2^{d+1} \sum_{j\geq 2} \frac{1}{[j]!} \sum_{h\geq 1} \frac{1}{(h+j)((h+j)^d - 2^d)}; \tag{2}$$

see (50) for an alternative expression. In particular, $\mu_1 = 1/3$ and $\mu_2 = 4\pi^2 - 39$; see [26, 36].

The phase change. Our first result says that when d increases, there is a change of nature for the limit distribution of X_n.

Theorem 1. (i) *If* $1 \leq d \leq 8$, *then*

$$\frac{X_n - \mu_d n}{\sigma_d \sqrt{n}} \xrightarrow{\mathscr{M}} N(0, 1),$$

where $\xrightarrow{\mathscr{M}}$ *denotes convergence of all moments and* $N(0, 1)$ *is the standard normal random variable (zero mean and unit variance). The constants* σ_d *are given in (52).*

(ii) *If* $d \geq 9$, *then the sequence of random variables* $(X_n - \mathbb{E}(X_n))/\sqrt{\mathbb{V}(X_n)}$ *does not converge to a fixed limit law.*

In the first case, convergence in distribution of $(X_n - \mu_d n)/\sqrt{\sigma_d^2 n}$ is also implied.

Why phase change? One key (analytic) reason why the limiting behavior of X_n changes its nature for $d \geq 9$ is because of the second order term in the asymptotic expansion of $\mathbb{E}(X_n)$

$$\mathbb{E}(X_n) = \mu_d n + G_1(\beta \log n) n^\alpha + o(n^\alpha + n^\varepsilon) \qquad (d \geq 2), \tag{3}$$

where $\alpha := 2\cos(2\pi/d) - 1$, $\beta := 2\sin(2\pi/d)$, and $G_1(x)$ is a bounded, 1-periodic function; see (49) for an explicit expression. If $d \leq 8$, then $\alpha < 1/2$; and $\alpha \in (1/2, 1)$ if $d \geq 9$; see Table 1 for numeric values of α.

d	2	3	4	5	6	7	8	9
α	-3	-2	-1	-0.38	0	0.24	0.41	0.53

Table 1: *Approximate numeric values of* $\alpha = 2\cos(2\pi/d) - 1$ *for* d *from* 2 *to* 9.

From this expansion, we can derive the asymptotics of the variance

$$\mathbb{V}(X_n) \sim \begin{cases} \sigma_d^2 n, & \text{if } 1 \le d \le 8; \\ G_2(\beta \log n) n^{2\alpha}, & \text{if } d \ge 9, \end{cases} \tag{4}$$

where $G_2(x)$ is a bounded, 1-periodic function.

Intuitively, we see that the periodicity in (3) becomes more pronounced as d grows (see Figure 2), implying larger and larger variance in (4), so that in the end $(X_n - \mathbb{E}(X_n))/\sqrt{\mathbb{V}(X_n)}$ does not converge to a fixed limit law.

Phase changes in other search trees. The situation here is similar to several phase change phenomena already studied in the literature in many varieties of random search trees and related algorithms: m-ary search trees, fringe-balanced binary search trees, generalized quicksort, etc; see [2, 3, 7, 15, 28, 29]. See also Janson [33] for a very complete description of phase changes in urn models, which are closely connected to many random search trees.

However, the analytic context here is much more involved than previously studied search trees because, as we will see, the underlying differential equation is no more of Cauchy-Euler type, which demands more delicate analysis.

Phase changes in random fragmentation models. The same phase change phenomenon as leaves in random quadtrees was first observed in Dean and Majumdar [8], where they proposed *random continuous fragmentation models* to *explain heuristically* the phase changes in random search trees. Their continuous model corresponding to quadtrees is as follows. Pick a point in $[0, x]^d$ uniformly at random ($x \gg 1$), which then splits the space into 2^d smaller hyperrectangles. Continue the same procedure in the sub-hyperrectangles whose volumes are larger than unity. The process stops when all sub-hyperrectangles have volumes less than unity. They argue heuristically that the total number of splittings undergoes a phase change: "While we can rigorously prove that the distribution is indeed Gaussian in the sub-critical regime [$d \le 8$], we have not been able to calculate the full distribution in the super-critical regime [$d \ge 9$]"; see [8].

Recently, Janson (private communication) showed that the same type of phase change can be constructed by considering the number of nodes at distance ℓ satisfying $\ell \bmod d \equiv j$, $0 \le j < d$, in random binary search trees, or equivalently, the number of nodes using the $(\ell+1)$-st coordinate as discriminators in random k-d trees, where $\ell \bmod d \equiv j$.

Recurrence. By the recursive nature of the problem proper, X_n satisfies the recurrence

$$X_n \stackrel{\mathscr{D}}{=} X_{J_1}^{(1)} + \cdots + X_{J_{2^d}}^{(2^d)} + \delta_{n,1} \qquad (n \ge 1), \tag{5}$$

with $X_0 = 0$, where the symbol $\stackrel{\mathscr{D}}{=}$ denotes equality in distribution, the J_i's and the $X_n^{(i)} \stackrel{\mathscr{D}}{=} X_n$'s are independent, $\delta_{n,1}$ denotes the Kronecker symbol, and

$$\begin{aligned} \pi_{n,\mathbf{j}} &:= \mathbb{P}(J_1 = j_1, \cdots J_{2^d} = j_{2^d}) \\ &= \binom{n-1}{j_1, \ldots, j_{2^d}} \int_{[0,1]^d} q_1(\mathbf{x})^{j_1} \cdots q_{2^d}(\mathbf{x})^{j_{2^d}} \, \mathrm{d}\mathbf{x}, \end{aligned}$$

denotes the probability that the 2^d subtrees of the root are of sizes $j_1, \dots, j_{2^d}$. Here $\mathrm{d}\mathbf{x} = \mathrm{d}x_1 \cdots \mathrm{d}x_d$ amd the $q_i(\mathbf{x})$'s denote the volumes of the hyperrectangles split by a random point $\mathbf{x} = (x_1, \dots, x_d)$. We can arrange the $q_i(\mathbf{x})$'s as follows

$$q_h(\mathbf{x}) = \prod_{1 \le i \le d} ((1-b_i)x_i + b_i(1-x_i)) \qquad (1 \le h \le 2^d), \tag{6}$$

where $(b_1, \dots, b_d)_2$ stands for the binary representation of $h-1$ (the first few digits being completed with zeros if $\lfloor \log_2(h-1) \rfloor < d-1$, so that $0 = (\underbrace{0, \dots, 0}_{d})_2$, $1 = (\underbrace{0, \dots, 0}_{d-1}, 1)_2$, etc.).

The moment-transfer approach. By (5), all moments of X_n (centered or not) satisfy the same recurrences of the form

$$A_n = B_n + 2^d \sum_{0 \le j < n} \pi_{n,j} A_j \qquad (n \ge 1), \tag{7}$$

with A_0 and $\{B_n\}_{n \ge 1}$ given, where

$$\pi_{n,j} = \binom{n-1}{j} \int_{[0,1]^d} (x_1 \cdots x_d)^j \, (1 - x_1 \cdots x_d)^{n-1-j} \, \mathrm{d}\mathbf{x}. \tag{8}$$

Many different expressions for $\pi_{n,j}$ can be found in [19, 34]; see also [25].

To prove the limit distribution, we apply the *moment-transfer approach*, which has proved successful in diverse problems of recursive nature. We have applied the approach to and developed the required asymptotic tools for many problems, including m-ary search trees, generalized quicksort and most variations of quicksort, bucket digital search trees, maximum-finding algorithms in distributed networks, maxima in right triangle; see the survey paper [29] for more references.

The basic idea of the approach is, because all moments satisfy the same recurrence (7), to incorporate the analysis of the asymptotics of higher moments into developing the so-called *asymptotic transfer*, which, roughly speaking, infers asymptotics of A_n from that of B_n. Such an approach always reduces most analysis to obtaining the first or second moments, the remaining part being more or less mechanical. It also offers the possibility of refining the limit theorems by stronger approximation results like convergence rates and local limit theorems, the new ingredients needed being developed in [28] for m-ary search trees; see also [1].

Second phase change. The refined moment-transfer approach (see [28]) shows that X_n undergoes a second phase change in convergence rate to normal limit law (often referred to as the Berry-Esseen bound). Our result says that the convergence rate to normal law is of order $n^{-1/2}$ when $1 \le d \le 7$, but is of a poorer order $n^{-3(3/2-\sqrt{2})} \approx n^{-0.24}$ when $d = 8$. Both rates are optimal modulo the implied constants. We will indeed derive local limit theorems for X_n, which are more precise and informative than convergence in distribution.

Resolution of the recurrence (7). *Exact solutions* of the recurrence (7) were first investigated by Flajolet et al. in [19] (see also [36, 39]), based mainly on the crucial introduction of the Euler transform. *Asymptotic properties* of (7) were also thoroughly examined in [19], using powerful complex-analytic tools. Their approach is very efficient in deriving the asymptotic expansions, but requires stronger information on the given "toll sequence" B_n.

In this paper, we show that the exact solution given via Euler transform in [19] (see (19)) can also be obtained by using the usual Poisson generating functions. Although this approach is essentially the same as the Euler transform on ordinary generating functions, it offers an operational advantage in simplifying the calculation of the exact variance; see Section 3.2.

Asymptotic transfer of the recurrence (7). We will develop the asymptotic transfer needed for deriving asymptotics of moments. Most proofs of previously known phase changes in random search trees and quicksort algorithms rely more or less on developing the asymptotic transfer for Cauchy-Euler differential equations (abbreviated as DEs) of the form

$$\text{Polynomial}(\vartheta)\xi(z) = \eta(z), \tag{9}$$

where η is independent of ξ and $\vartheta := (1-z)(\mathrm{d}/\mathrm{d}z)$. The main transfer problem under this framework is to derive asymptotics of $[z^n]\xi(z)$ when that of $[z^n]\eta(z)$ is known, where $[z^n]\xi(z)$ denotes the coefficient of z^n in the Taylor expansion of f. A very general, elementary asymptotic theory for such DEs with a large number of applications is given in [7], the origin of such a development being traceable to Sedgewick's analysis on quicksort (see [45]).

For quadtrees, the DE satisfied by the generating function $A(z) := \sum_n A_n z^n$ is given by

$$\vartheta(z\vartheta)^{d-1}(A(z) - B(z)) = 2^d A(z), \tag{10}$$

which is not of the type (9) but can be rewritten in the extended form

$$P_0(\vartheta)A(z) = \vartheta(z\vartheta)^{d-1}B(z) + \sum_{1\le j<d}(1-z)^j P_j(\vartheta)A(z), \tag{11}$$

where $P_0(x) = x^d - 2^d$ and the $P_j(x)$'s are polynomials of degree d; see (23).

We then extend the iterative operator approach introduced in [5] to analyzing the expected cost of partial match queries in random k-d trees. The approach turns out to be very useful for extended Cauchy-Euler DEs of the form (11); see [6] for another application to consecutive records in random sequences.

The main differences of the current application from the previous ones are: (i) we consider general non-homogeneous part (or toll functions) rather than specific ones; (ii) the method of Frobenius (and the method of annihilators) used in our previous papers is avoided and replaced by a more uniform elementary argument, the resulting proof being completely elementary and requiring almost no knowledge on DE; (iii) we give not only necessary but also sufficient conditions for all transfers we developed; the same proof for the sufficiency part also easily modified for proving the necessity in all cases, keeping uniformity of the approach; (iv) the proof we give in its current form is easily amended for more general DEs with polynomial coefficients; (v) we put forth means of simplifying the expressions for the constants involved; the resulting expressions are in some cases simpler than those derived in [19]; also our expressions are easily amended for numeric purposes.

A universal condition for asymptotic linearity? One main result our approach can achieve states that A_n is asymptotically linear $A_n \sim Kn$ if and only if $B_n = o(n)$ and the series $\sum_n B_n n^{-2}$ is convergent, where K is explicitly given in terms of the B_n's; see (16). It is interesting to see that exactly the same condition for the asymptotic linearity of A_n holds for other recurrences appearing in quicksort, m-ary search trees, generalized quicksort, and many others; see [7]. Note that the expression for the linearity constant K differs from one case to another. The series condition $|\sum_n B_n n^{-2}| < \infty$ also arises in many other problems such as generalized subadditive inequalities, divide-and-conquer algorithms, large deviations, etc.; see [31] and the references therein. Is there a deeper reason why the series condition is so universal?

Organization of the paper. In the next Section 2, we develop general asymptotic transfer results, which can be applied to more general shape characteristics and cost measures. In Sections 3 and 4, we study the phase change phenomena exhibited by the number of leaves and discuss the extension to general cost measures. Effective numerical procedures will also be given of computing the limiting mean and variance constants for X_n. The extension of our consideration to Devroye's grid-trees (see [12]) is given in the final section.

Notation. Throughout this paper, the notation $[z^n]f(z)$ denotes the coefficient of z^n in the Taylor expansion of f. The generic symbol ε always represents some small quantity whose value may vary from one occurrence to another; similarly, the generic symbol c stands for a suitable constant. We define two operators $\mathbb{D}_z := \mathrm{d}/\mathrm{d}z$ and $\vartheta := (1-z)\mathbb{D}_z$. The same set of symbols $\{B_n, B(z), B^*(s)\}$ is used for the sequence B_n, its generating function $B(z) = \sum_n B_n z^n$, and its factorial series or Mellin transform $B^*(s) = \int_0^1 (1-x)^{s-1} B(x)\,\mathrm{d}x$, respectively.

2 Asymptotic transfer of the quadtree recurrence

We develop the asymptotic tools in this section by proving the different types of asymptotic transfer needed for later uses. A salient feature of our transfers is that the asymptotic condition in each case is not only sufficient but also proved to be necessary.

Three types of asymptotic transfer. For simplicity, we assume $A_0 = 0$ since otherwise the difference is given explicitly by $A_0(2^d-1)n + A_0$; see (19).

Theorem 2. *Let A_n be defined by the recurrence (7) with A_0 and $\{B_n\}_{n\ge1}$ given. Then*

(i) *(Small toll functions)*

$$A_n \sim K_B n \quad \textit{iff} \quad B_n = o(n) \ \ \textit{and} \ \ \left|\sum_n B_n n^{-2}\right| < \infty, \tag{12}$$

where the constant K_B is given in (16);

(ii) *(Linear toll functions) Assume that $B_n = cn + u_n$, where $c \in \mathbb{C}$ and u_n is a sequence of complex numbers. Then*

$$A_n \sim \frac{2}{d} cn\log n + K_1 n \quad \textit{iff} \quad u_n = o(n) \ \textit{and} \left|\sum_n u_n n^{-2}\right| < \infty, \tag{13}$$

where $K_1 := cK_2 + K_u$ with K_u defined by replacing the sequence B_n by u_n in (16) and K_2 given explicitly by

$$K_2 := -1 - \frac{2}{d} + 2\gamma + \frac{2}{d}\sum_{1\le j<d} \psi(2 - 2e^{2j\pi i/d}), \tag{14}$$

ψ being the logarithmic derivative of the Gamma function (see [14]);

(iii) *(Large toll functions) Assume that $\Re(\upsilon) > 1$ and $c \in \mathbb{C}$. Then*

$$B_n \sim cn^{\upsilon} \quad \textit{iff} \quad A_n \sim \frac{c(\upsilon+1)^d}{(\upsilon+1)^d - 2^d} n^{\upsilon}. \tag{15}$$

More refinements to (12) under stronger assumptions on B_n will be proved below.

The linearity constant. Given a sequence B_n, define the constant K_B by the series

$$K_B = \frac{2}{d}\sum_{k\geq 0} V_k B^*(k+2), \tag{16}$$

which is absolutely convergent under the condition (12) on B_n, where V_k is defined recursively by $V_k = 0$ when $k < 0$, $V_0 = 1$, and

$$V_k = \sum_{1\leq \ell < d} \frac{P_\ell(k+2)}{P_0(k+2)} V_{k-\ell} \qquad (k \geq 1), \tag{17}$$

and the function B^* is given by

$$B^*(s) := \int_0^1 B(x)(1-x)^{s-1}\,\mathrm{d}x = \sum_{j\geq 1} \frac{B_j j!}{s(s+1)\cdots(s+j)}, \tag{18}$$

when the integral and series converge. Here the polynomials $P_j(x)$'s are given in (23). Note that when $d = 1$, $V_k = \delta_{k,0}$, so that $K_B = 2B^*(2)$; see [30].

2.1 Euler transform and Poissonization

Euler transform. Flajolet et al. proposed in [19] an approach via Euler transform for solving the recurrence (7); their result is

$$A_n = A_0 + n\left((2^d-1)A_0 + B_1\right) + \sum_{2\leq k\leq n}\binom{n}{k}(-1)^k \sum_{2\leq j\leq k}\left(B_j^\star - B_{j-1}^\star\right)\prod_{j<\ell\leq k}\left(1-\frac{2^d}{\ell^d}\right), \tag{19}$$

for $n \geq 0$, where $B_n^\star$ denotes the Euler transform of the sequence B_n

$$B_n^\star := \sum_{1\leq j\leq n}\binom{n}{j}(-1)^j B_j.$$

As one can see from (19), the appearance of $B_n^\star$ and the power of -1 makes the asymptotics of A_n less transparent.

Poissonization. An alternative way of deriving (19) is as follows. Consider the Poisson generating functions of both sequences: $\tilde{A}(z) := e^{-z}\sum_{n\geq 0} A_n z^n/n!$ and $\tilde{B}(z) := e^{-z}\sum_{n\geq 1} B_n z^n/n!$. Then (7) translates into

$$\tilde{A}'(z) + \tilde{A}(z) = \tilde{B}'(z) + \tilde{B}(z) + 2^d \int_{[0,1]^d} \tilde{A}(x_1\cdots x_d z)\,\mathrm{d}\mathbf{x},$$

with the initial condition $\tilde{A}(0) = A_0$. Let $\tilde{A}_n := n![z^n]\tilde{A}(z)$ and $\tilde{B}_n := n![z^n]\tilde{B}(z)$. Then

$$\tilde{A}_n + \tilde{A}_{n-1} = \tilde{B}_n + \tilde{B}_{n-1} + \frac{2^d}{n^d}\tilde{A}_{n-1} \qquad (n \geq 1), \tag{20}$$

(for convenience, defining $B_0 = \tilde{B}_0 = 0$). Observe that

$$\tilde{A}_n = (-1)^n A_n^\star = (-1)^n \sum_{0\leq k\leq n}\binom{n}{k}(-1)^k A_k,$$

and $\tilde{B}_n = (-1)^n B_n^\star$. By iterating the recurrence (20) and by taking into account the initial values, we obtain (19).

Although the approach is essentially the same as that via Euler transform, it is helpful in deriving a dimension-free expression for, say the variance of X_n; see Section 3.2. It also offers the possibility of obtaining the asymptotics of A_n by the usual Mellin transform techniques.

Asymptotics of the recurrence (7). A very powerful complex-analytic approach is proposed in [19] to the asymptotics of (7). The main idea is to apply singularity analysis (see [21]); so one needs the asymptotics of the generating function $\sum_n A_n z^n$ for $z \sim 1$, which, by the Euler transform, leads to the study of the generating function $A^\star(t) := \sum_n A_n^\star t^n$ for t near $-\infty$. For that purpose, they apply integral representation for $A^\star(-t)$ of the form

$$A^\star(-t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\pi t^s}{\sin \pi s}\, \varphi(s)\, \mathrm{d}s,$$

for suitably chosen c and $\varphi(s)$ satisfying $\varphi(k) = A_k^\star$ for $k \geq 2$. The determination of such an "analytic extrapolation" of $A_k^\star$ to complex s is crucial.

The major limitation of this approach is that when the given sequence B_n is, say only known up to $O(n^\alpha)$ or $\sim n^\alpha$ for some α, it is not obvious how to find an analytic extrapolation and then to deduce the right order of A_n because of the underlying "exponential cancellations of order": roughly, $\binom{n}{k}$ has its largest term of order $2^n n^{-1/2}$, but most of our sequences grow only polynomially in n; see [23] for asymptotics on alternating binomial sums.

Alternatively, one might try the usual Mellin analysis for $\tilde{A}(z)$ (or its truncated functions); again analytic properties of the involved function at $\sigma \pm i\infty$ may be very challenging.

Note that the value A_0 and the sequence $\{B_n\}_{n\geq 1}$ are enough to completely determine the sequence A_n. This property will be useful in our numeric procedure; see Section 3.2.

2.2 Asymptotic transfer I. Small toll functions

We prove the first case of Theorem 2 in this section by extending the approach we proposed before for the analysis of k-d trees. The main idea is to write the underlying DE in the form of certain "perturbed" DE of Cauchy-Euler type, and then to use some iterative operator arguments.

The DE. Let $A(z) = \sum_{n\geq 0} A_n z^n$ and $B(z) = \sum_{n\geq 1} B_n z^n$. Then the recurrence (7) translates into the DE (10), which becomes simpler by considering $f := A - B$:

$$\left(\vartheta(z\vartheta)^{d-1} - 2^d\right) f(z) = 2^d B(z). \tag{21}$$

This DE can be re-written as the "perturbed" Cauchy-Euler DE

$$\begin{cases} P_0(\vartheta) f(z) = g(z) + 2^d B(z); \\ g(z) := \sum_{1\leq j<d} (1-z)^j P_j(\vartheta) f(z), \end{cases} \tag{22}$$

where $P_0(x) = x^d - 2^d$, and by induction

$$P_j(x) = (-1)^{j-1}[z^{d-1-j}] \prod_{0\leq r\leq j} \frac{x-r}{1-z(x-r)} \qquad (1 \leq j < d). \tag{23}$$

Note that all P_j's are polynomials of degree d; they can also be computed recursively as follows. Write

$$\vartheta(z\vartheta)^{d-1}f(z) = \sum_{0\le j<d}(1-z)^j\tilde{P}_{d,j}(\vartheta)f(z).$$

Then $P_j(x) = -\tilde{P}_{d,j}(x)$ for $1 \le j < d$. Here $\tilde{P}_{d,j}(x) = (x-j)(\tilde{P}_{d-1,j}(x) - \tilde{P}_{d-1,j-1}(x))$ with the boundary conditions $\tilde{P}_{1,0}(x) = x$, $\tilde{P}_{d,j}(x) = 0$ if $j < 0$ or $j \ge d$.

Let λ_j's denote the zeroes of $P_0(x) = 0$, namely, $\lambda_j = 2e^{2j\pi i/d}$ for $0 \le j < d$. In particular, $\lambda_0 = 2$.

All initial conditions zero. For convenience, we assume temporarily that all initial values are zeros $f^{(j)}(0) = 0$ for $0 \le j < d$. This implies that $\vartheta^j f(0) = 0$ for $0 \le j < d$ since

$$\vartheta^j f(z) = \sum_{0\le\ell\le j}(-1)^{j+\ell}S(j,\ell)(1-z)^\ell f^{(\ell)}(z),$$

where $S(j,\ell)$ represents the Stirling numbers of the second kind.

The Cauchy-Euler solution. Regarding the DE (22) as a Cauchy-Euler DE, we can then decompose the DE as follows.

$$(\vartheta-\lambda_{d-1})\cdots(\vartheta-\lambda_1)(\vartheta-2)f(z) = g(z) + 2^d B(z), \tag{24}$$

whose solution (exact or asymptotic) can be obtained by successively solving the first-order DE of the form

$$(\vartheta - \upsilon)\xi(z) = \eta(z),$$

which is given by

$$\xi(z) = \xi(0)(1-z)^{-\upsilon} + (1-z)^{-\upsilon}\int_0^z (1-t)^{\upsilon-1}\eta(t)\,\mathrm{d}t,$$

in the sense of formal power series; see [7].

Since all initial conditions are zero, we thus obtain the solution

$$f(z) = \left(\mathbf{I}_{\lambda_{d-1}} \circ \cdots \circ \mathbf{I}_{\lambda_1} \circ \mathbf{I}_2\right)[g + 2^d B](z), \tag{25}$$

where

$$\mathbf{I}_\upsilon[\phi](z) = (1-z)^{-\upsilon}\int_0^z (1-x)^{\upsilon-1}\phi(x)\,\mathrm{d}x. \tag{26}$$

Note that the function g involves itself f.

Thus the next steps consist of (i) clarifying the changes in asymptotic approximation under consecutive applications of the linear operators, and (ii) simplifying the resulting leading constants.

Asymptotic transfer for the linear operator.

Lemma 1 ([7]). (i) *(Small toll functions) Let $\upsilon \in \mathbb{C}$. If $\int_0^1 (1-x)^{\upsilon-1}\phi(x)\,\mathrm{d}x$ converges, then*

$$[z^n]\mathbf{I}_\upsilon[\phi](z) \sim \frac{n^{\upsilon-1}}{\Gamma(\upsilon)}\int_0^1 (1-x)^{\upsilon-1}\phi(x)\,\mathrm{d}x, \tag{27}$$

where Γ denotes the Gamma function.

(ii) *(Large toll functions) Let $\upsilon \in \mathbb{C}$. If $[z^n]\phi(z) \sim cn^\tau$, where $c \in \mathbb{C}$ and $\Re(\tau) > \Re(\upsilon) - 1$, then*

$$[z^n]\,\mathbf{I}_\upsilon[\phi](z) \sim \frac{c}{\tau + 1 - \upsilon} n^\tau. \tag{28}$$

Note that if $\upsilon = 0, -1, \ldots$ in case (i), then the $\sim$-transfer (27) becomes an o-transfer; similarly, if $c = 0$ in case (ii), then (28) becomes an o-transfer.
Proof. (Sketch) The estimate (27) follows from (26), and (28) from the expression

$$[z^n]\,\mathbf{I}_\upsilon[\phi](z) = \frac{\Gamma(n+\upsilon)}{\Gamma(n+1)} \sum_{0 \le k < n} \frac{\Gamma(k+1)}{\Gamma(k+1+\upsilon)} [z^k]\phi(z); \tag{29}$$

see [7]. ∎

Asymptotic linearity. We now prove the small toll functions part of Theorem 2 when $B_n = o(n)$ and $\sum_n B_n n^{-2}$ converges. The assumption that the series $\sum_n B_n n^{-2}$ converges implies that $|\int_0^1 (1 - x)B(x)\,\mathrm{d}x| < \infty$. Assume at the moment that

$$\left| \int_0^1 (1-x) g(x)\,\mathrm{d}x \right| < \infty. \tag{30}$$

Then by applying consecutively Lemma 1, we obtain

$$A_n = [z^n] f(z) + B_n = \frac{K'}{P_0'(2)} n + o(n), \tag{31}$$

where

$$K' := \int_0^1 (1-x) \left(g(x) + 2^d B(x) \right)\,\mathrm{d}x = \sum_{j \ge 0} \frac{[z^j]g(z) + 2^d B_j}{(j+1)(j+2)}. \tag{32}$$

The next step is to prove (30).

Proof of (30). Define

$$\Lambda(s) := \int_0^1 (1-x)^{s-1} P_0(\vartheta) f(x)\,\mathrm{d}x,$$

where the ϑ-operator is understood to be $(1-x)\mathrm{d}/\mathrm{d}x$.

Since $B_n = o(n) = o(n^{1+\varepsilon})$, $A_n = o(n^{1+\varepsilon})$ by (46) below. Thus $f(x) = O((1-x)^{-2-\varepsilon})$ for $0 \le x < 1$ and

$$P_0(\vartheta) f(x) = O(f^{(d)}(x)) = O((1-x)^{-d-2-\varepsilon}),$$

for $0 \le x < 1$. It follows that $\Lambda(s)$ is finite for sufficiently large s, say $s \ge s_0 \ge d + 2 + \varepsilon$. We show that we can take $s_0 = 2$. Note that $\Lambda(s)$ is an analytic function in the half-plane $\Re(s) \ge 2$, but for our purposes we need only real values of s.

Lemma 2 ([5]). *Let $p(x)$ and $q(x)$ be two polynomials of degrees at most d. Assume that $\phi(x)$ is defined in the unit interval with $\phi^{(j)}(0) = 0$ for $0 \le j < k$. Then*

$$\int_0^1 (1-x)^{s-1} \left(p(\vartheta) q(\vartheta)^{-1} \right) \phi(x)\,\mathrm{d}x = \frac{p(s)}{q(s)} \int_0^1 (1-x)^{s-1} \phi(x)\,\mathrm{d}x, \tag{33}$$

provided that $q(s) \ne 0$ and that both integrals converge.

Substituting (22) into the integral and applying (33), we see that $\Lambda(s)$ satisfies the difference equation

$$\Lambda(s) = 2^d B^*(s) + \sum_{1\le j<d} \frac{P_j(j+s)}{P_0(j+s)} \Lambda(j+s). \tag{34}$$

By assumption, $B^*(s)$ is finite for $s \ge 2$. Also $\Lambda(s)$ is bounded for $s \ge d+2+\varepsilon$ as showed above. Thus by iterating the equation (34), we deduce that $\Lambda(s)$ is finite for $s \ge 2$.

This proves (30) because

$$\int_0^1 (1-x)g(x)\,\mathrm{d}x = \int_0^1 (1-x)\left(P_0(\vartheta)f(x) - 2^d B(x)\right)\,\mathrm{d}x,$$

and from (32), it follows that $K' = \Lambda(2)$.

Further simplification of the constant K'. Taking first $s = 2$ in (34) and then iterating the recurrence (34) N times, we get

$$K' = K'_N + \sum_{1\le j\le N(d-1)+1} \frac{e_{N,j}}{P_0(j+N+1)} \Lambda(j+N+1),$$

where $e_{1,j} = P_j(j+2)$ for $1 \le j \le d$,

$$e_{N,j} := \sum_{1\le \ell\le d} \frac{P_\ell(j+N+1)}{P_0(j+N+1-\ell)}\, e_{N-1,j+1-\ell} \qquad (1 \le j \le N(d-1)+1),$$

for $N \ge 2$, and

$$K'_N = 2^d \left(B^*(2) + \sum_{1\le j\le (N-1)d} \frac{B^*(j+2)}{P_0(j+2)} \sum_{1\le\ell\le j} e_{\ell,j+1-\ell} \right),$$

for $N \ge 0$.

Since $\Lambda(N) \to 0$ as $N \to \infty$, we have

$$K' = \lim_{N\to\infty} K'_N = 2^d \left(B^*(2) + \sum_{j\ge 1} \frac{B^*(j+2)}{P_0(j+2)} \sum_{1\le\ell\le j} e_{\ell,j+1-\ell} \right).$$

Define

$$V_k := \frac{1}{P_0(k+2)} \sum_{1\le\ell\le k} e_{\ell,k+1-\ell}.$$

Then V_k satisfies (17) and we have

$$K' = 2^d \sum_{k\ge 0} B^*(k+2) V_k.$$

It follows, by (31), that $K_B = K'/P'_0(2)$.

Absolute convergence of the series representation (16) for K_B. There is no *a priori* reason that the series representation for K_B in (16) is convergent. We show that under the assumptions on B_n in (12) the series in (16) is indeed absolutely convergent.

Observe first that by the factorial series expression in (18)

$$B^*(k+2) = O(k^{-2}).$$

We need then an estimate for V_k.

If $d = 2$, then $P_1(s) = s(s-1)$, and we can solve the recurrence of V_k explicitly, giving

$$V_k = 12\frac{k+1}{(k+3)(k+4)} \qquad (k \geq 0). \tag{35}$$

Consequently,

$$\begin{aligned} K_B &= 12\sum_{k\geq 0} \frac{k+1}{(k+3)(k+4)} B^*(k+2) \\ &= 12\int_0^1 B(x)\left(\frac{1+2x}{(1-x)^3}\log\frac{1}{x} - \frac{5+x}{2(1-x)^2}\right) \mathrm{d}x; \end{aligned}$$

see also [36, 39].

Lemma 3. *The sequence V_k satisfies the estimate*

$$V_k = O\left(k^{-1}(\log k)^{d-2}\right), \tag{36}$$

for $d \geq 2$.

The order is tight; indeed, we can derive a more precise asymptotic approximation; see (39) below.

Proof. We first show that the generating function $V(z)$ of V_k satisfies the DE

$$\mathbb{D}_z\left(z(1-z)\mathbb{D}_z\right)^{d-1}\left(z^2V(z)\right) - 2^d zV(z) = 0. \tag{37}$$

By Cauchy's integral representation for V_k

$$V_k = \frac{1}{2\pi i}\oint_{|w|=\varepsilon} w^{-k-1}V(w)\,\mathrm{d}w = \frac{1}{2\pi i}\oint_{|w-1|=\varepsilon} (1-w)^{-k-1}V(1-w)\,\mathrm{d}w.$$

Then, by the relation (see (17)),

$$P_0(k+2)V_k - \sum_{1\leq \ell<d} P_\ell(k+2)V_{k-\ell} = 0,$$

we have

$$\begin{aligned} 0 &= \frac{1}{2\pi i}\oint_{|w-1|=\varepsilon}(1-w)V(1-w)\left[P_0(k+2)(1-w)^{-k-2} - \sum_{1\leq\ell<d} P_\ell(k+2)(1-w)^{-k+\ell-2}\right]\mathrm{d}w \\ &= \frac{1}{2\pi i}\oint_{|w-1|=\varepsilon}(1-w)V(1-w)\left[\vartheta_w(w\vartheta_w)^{d-1} - 2^d\right](1-w)^{-k-2}\,\mathrm{d}w, \end{aligned}$$

by the definition of the P_j's, where $\vartheta_w := (1-w)\mathrm{d}/\mathrm{d}w$. It follows, by multiplying both sides by z^k and then summing over all nonnegative k, that

$$I_d(z) - 2^d V(z) = 0,$$

where

$$I_d(z) := \frac{1}{2\pi i}\oint_{|w-1|=\varepsilon} (1-w)V(1-w)\left[\vartheta_w(w\vartheta_w)^{d-1}\right]\frac{(1-w)^{-2}}{1-\frac{z}{1-w}}\,\mathrm{d}w.$$

By successive integration by parts, we have

$$\begin{aligned} I_d(z) &= \frac{(-1)^d}{2\pi i}\oint_{|w-1|=\varepsilon}\frac{(1-w)^{-2}}{1-\frac{z}{1-w}}\mathbb{D}_w\left(w(1-w)\mathbb{D}_w\right)^{d-1}\left((1-w)^2V(1-w)\right)\,\mathrm{d}w \\ &= \frac{1}{2\pi i}\oint_{|w|=\varepsilon}\frac{w^{-2}}{1-\frac{z}{w}}\mathbb{D}_w\left(w(1-w)\mathbb{D}_w\right)^{d-1}\left(w^2V(w)\right)\,\mathrm{d}w, \end{aligned}$$

where $\mathbb{D}_w := \mathrm{d}/\mathrm{d}w$. This proves (37).

By Frobenius method (see [32]), we seek solutions of the form $V(z) = (1-z)^{-s}\xi(1-z)$ with ξ analytic at zero. Substituting such a form into (37) gives for $d = 1$

$$I_1(z) \sim \xi(0)s(1-z)^{-s-1} \qquad (z\sim 1).$$

By induction, we obtain

$$I_d(z) \sim \xi(0)s^d(1-z)^{-s-1} \qquad (z\sim 1).$$

Thus, the indicial equation is $s^d = 0$, implying that

$$V(z) = O\left(\log^{d-1}|1-z|\right) \qquad (z\sim 1).$$

It follows, by singularity analysis (see [21]), that V_k satisfies the estimate (36). This proves Lemma 3. ∎

A more precise approximations to the asymptotics of V_k. Since the generating function of the sequence V_k satisfies the explicit, homogeneous DE (37), we can derive more precise asymptotic estimates as follows.

By applying either the Euler transform approach of [19] or the Poisson generating functions, we obtain

$$V_k = \sum_{1\le \ell\le k+1}\binom{k+1}{\ell}(-1)^{\ell+1}\ell\prod_{1\le j<d}\frac{\Gamma(3-\lambda_j)\Gamma(\ell+1)}{\Gamma(\ell+2-\lambda_j)} \qquad (k\ge 0).$$

Consequently, we have the integral representation (see [23])

$$V_k = \frac{1}{2\pi i}\int_{\varepsilon-i\infty}^{\varepsilon+i\infty}\frac{\Gamma(k+2)\Gamma(1-s)}{\Gamma(k+2-s)}\prod_{1\le j<d}\frac{\Gamma(3-\lambda_j)\Gamma(s+1)}{\Gamma(s+2-\lambda_j)}\,\mathrm{d}s. \tag{38}$$

From this representation, we can show that

$$V_k \sim \frac{d2^{d-1}(2^d-1)}{(d-2)!}k^{-1}(\log k)^{d-2}, \tag{39}$$

for $d \geq 2$ and large k. Note that the leading constants first grows and then decreases to zero

$$\left\{\frac{d2^{d-1}(2^d-1)}{(d-2)!}\right\}_{d\geq 2} = \left\{12, 84, 240, 413\tfrac{1}{3}, 504, 474\tfrac{2}{15}, 362\tfrac{2}{3}, 233\tfrac{3}{5}, 129\tfrac{19}{21}, 63\tfrac{1531}{2835}, \cdots\right\}.$$

Since the leading constants are quite large for small d, the convergence of the series (16) is poor for small d; we will propose a more efficient numeric procedure for computing K_B.

In particular, if $d = 2$, the integrand has three simple poles at $s = -1, -2$, and -3, and the residues of these poles add up to $12(k+1)/((k+3)(k+4))$, in accordance with (35). But for $d \geq 3$, the resulting expressions are more complicated because there are infinitely many poles.

An integral representation for the constant K_B. By substituting the expression (38) of V_k in (16), we obtain

$$K_B = \frac{2}{2d\pi i}\int_{c-i\infty}^{c+i\infty} \Upsilon(s) \prod_{1\leq j<d} \frac{\Gamma(3-\lambda_j)\Gamma(s+1)}{\Gamma(s+2-\lambda_j)}\,\mathrm{d}s, \tag{40}$$

where

$$\Upsilon(s) := \sum_{k\geq 0} B^*(k+2)\frac{\Gamma(k+2)\Gamma(1-s)}{\Gamma(k+2-s)},$$

and $c > -1$ lies in the half-plane where the series on the right-hand side converges. Thus if analytic properties of Υ are known, then K_B can be further simplified; see for example (44). Also if $d = 2$, then $K_B = 12(\Upsilon(-1) - 2\Upsilon(-2) + \Upsilon(-3))$; see (35).

Nonzero initial conditions. We now prove that the linearity constant K_B is of the form (16) even with nonzero initial conditions.

We start from making all the initial conditions zero

$$\bar{f}(z) := f(z) - \sum_{0\leq j<d} (A_j - B_j)z^j,$$

so that, by (21),

$$\left(\vartheta(z\vartheta)^{d-1} - 2^d\right)\bar{f}(z) = 2^d B(z) + 2^d C(z),$$

where (for convenience, defining $B_0 = 0$)

$$C(z) := \sum_{0\leq j<d} (A_j - B_j)\, z^j - 2^{-d}\left(\vartheta(z\vartheta)^{d-1}\right)\left(\sum_{0\leq j<d}(A_j - B_j)z^j\right).$$

By the same approach as above, we obtain $A_n \sim \bar{K}n$, where the linearity constant $\bar{K}$ is given by

$$\bar{K} = \frac{2}{d}\sum_{k\geq 0} V_k B^*(k+2) + \frac{2}{d}\sum_{k\geq 0} V_k \int_0^1 (1-x)^{k+1} \sum_{0\leq j<d}(A_j - B_j)x^j\,\mathrm{d}x + \bar{c}.$$

Here

$$\begin{aligned}\bar{c} &:= -\frac{2^{1-d}}{d}\sum_{k\geq 0} V_k \int_0^1 (1-x)^{k+1}\left(\vartheta_x(x\vartheta_x)^{d-1}\right)\left(\sum_{0\leq j<d}(A_j - B_j)x^j\right)\mathrm{d}x \\ &= -\frac{2^{1-d}}{d}\int_0^1 (1-x)V(1-x)\left(\vartheta_x(x\vartheta_x)^{d-1}\right)\left(\sum_{0\leq j<d}(A_j - B_j)x^j\right)\mathrm{d}x.\end{aligned}$$

By the same argument used to derive the DE satisfied by $V(z)$, we have

$$\bar{c} = -\frac{2^{1-d}}{d}\int_0^1 \left(\sum_{0\le j<d}(A_j-B_j)(1-x)^j\right)\mathbb{D}_x\left(x(1-x)\mathbb{D}_x\right)^{d-1}\left(x^2V(x)\right)\,\mathrm{d}x.$$

But by (37)

$$\mathbb{D}_x\left(x(1-x)\mathbb{D}_x\right)^{d-1}\left(x^2V(x)\right) = 2^d xV(x);$$

it follows that

$$\bar{c} = -\frac{2}{d}\int_0^1 (1-x)V(1-x)\left(\sum_{0\le j<d}(A_j-B_j)x^j\right)\mathrm{d}x.$$

Thus

$$\bar{K} = \frac{2}{d}\sum_{k\ge 0} V_k B^*(k+2);$$

this proves that the linearity constant is of the same form (16), which amounts to saying that *we do not need to nullify the initial conditions.*

An efficient numeric procedure. The above proof suggests a useful numeric procedure for computing the constant K_B. The crucial observation is that the first d terms we choose to be subtracted from $\bar{f}$ play no special role in our proof, meaning that we can indeed subtract a sufficiently large number, say N, of initial terms from f, resulting in a series form for K_B with convergence rate $(\log k)^{d-2}k^{-N}$. This is because the right-hand side of the DE is of order z^{N-1}, which yields, after taking the finite Mellin transform, the order k^{-N} for large k. Such a procedure quickly leads to a good numeric approximation to the leading constant K_B to high precision. We will apply this procedure to the constants appearing in the mean and variance of the number of leaves in Section 3.2

Necessity in (12). Assume that $A_n \sim cn$ for some constant c. The special form (8) or the following one (see [19])

$$\pi_{n,j} = \frac{1}{(d-1)!}\binom{n-1}{j}\int_0^1 (-\log t)^{d-1}t^j(1-t)^{n-1-j}\,\mathrm{d}t,$$

can be used to prove that $B_n = o(n)$ by (7). We propose instead a proof based again on linear operators, the advantage being generally applicable to more complicated recurrences while keeping uniformity of the proof.

By (21)

$$\begin{aligned} B(z) &= A(z) - 2^d\left(\vartheta^{-1}\left(z^{-1}\vartheta^{-1}\right)^{d-1}\right)A(z)\\ &= A(z) - 2^d\left(\mathbf{I}_0\circ\left(z^{-1}\mathbf{I}_0\right)^{d-1}\right)[A](z). \end{aligned}$$

Since $A_n \sim cn$, we have, by (28),

$$[z^n]\mathbf{I}_0[A](z) \sim \frac{c}{2}n, \quad [z^n]z^{-1}\mathbf{I}_0[A](z) \sim \frac{c}{2}n.$$

Applying successively these estimates yields

$$[z^n]2^d\left(\mathbf{I}_0\circ\left(z^{-1}\mathbf{I}_0\right)^{d-1}\right)[A](z) \sim cn.$$

Thus $B_n = o(n)$.

We then prove that $|\sum_n B_n n^{-2}| < \infty$ by showing that $B^*(2)$ is finite. By (34), it suffices to show that $\Lambda(2)$ is finite. Since $A_n \sim cn$ and $B_n = o(n)$, we deduce that $f(x) = O((1-x)^{-2})$ for $0 \le x < 1$. It follows that

$$\begin{aligned}\Lambda(2) &= \lim_{s\to 2^+} \Lambda(s)\\ &= \lim_{s\to 2^+} P_0(s)\int_0^1 (1-x)^{s-1} f(x)\,\mathrm{d}x\\ &= O(1).\end{aligned}$$

This complete the proof of (12).

2.3 Asymptotic transfer II. Linear toll functions

We prove part (ii) of Theorem 2 in this section. By the result of part (i), it suffices to consider the case when $B_n \equiv n$ for $n \ge 1$. Then $B(z) = z/(1-z)^2$.

All initial conditions zero. It is simpler, as in part (i), to consider

$$\bar{f}(z) := A(z) - B(z) - \sum_{0\le j<d} (A_j - B_j)z^j,$$

so that $\bar{f}$ satisfies the DE

$$\left(\vartheta(z\vartheta)^{d-1} - 2^d\right)\bar{f}(z) = 2^d B(z) + 2^d C(z),$$

with zero initial conditions, where

$$C(z) := \left(2^{-d}\vartheta(z\vartheta)^{d-1} - 1\right)\sum_{1\le j<d} (A_j - B_j)z^j.$$

Then $\bar{f}$ satisfies the DE

$$P_0(\vartheta)\bar{f}(z) = 2^d B(z) + 2^d C(z) + g(z),$$

where g is defined in (22), and for $n \ge d$

$$\begin{aligned}A_n &= [z^n]\left(\bar{f}(z) + B(z)\right)\\ &= n + [z^n]\left(\mathbf{I}_{\lambda_{d-1}}\cdots\circ\mathbf{I}_{\lambda_1}\circ\mathbf{I}_2\right)\left[2^d B + 2^d C + g\right](z).\end{aligned}$$

An expression for the iterates of the I-operators. Observe first that by integration by parts

$$(\mathbf{I}_\upsilon\circ\mathbf{I}_\tau)[\xi](z) = \frac{1}{\tau-\upsilon}\mathbf{I}_\tau[\xi](z) - \frac{1}{\tau-\upsilon}\mathbf{I}_\upsilon[\xi](z) \qquad (\upsilon\ne\tau),$$

so that by induction

$$\left(\mathbf{I}_{\lambda_{d-1}}\circ\cdots\circ\mathbf{I}_{\lambda_0}\right)[\xi](z) = \sum_{0\le j<d}\frac{\mathbf{I}_{\lambda_j}[\xi](z)}{\prod_{\ell\ne j}(\lambda_j-\lambda_\ell)}. \qquad (41)$$

Thus

$$\bar{f}(z) = \sum_{0\le j<d}\frac{\mathbf{I}_{\lambda_j}[2^d B + 2^d C + g](z)}{P_0'(\lambda_j)}.$$

The contribution of $2^d B(z)$. By applying (41), we have

$$\begin{aligned}
[z^n]\left(\mathbf{I}_{\lambda_{d-1}}\circ\cdots\circ\mathbf{I}_{\lambda_1}\circ\mathbf{I}_2\right)[2^dB](z) &= \sum_{0\le j<d}\frac{2^d}{P_0'(\lambda_j)}[z^n]\mathbf{I}_{\lambda_j}[B](z)\\
&= \frac{2^d}{P_0'(2)}[z^n]\left((1-z)^{-2}\log\frac{1}{1-z}-(1-z)^{-2}\right)\\
&\quad+\sum_{1\le j<d}\frac{2^d}{(2-\lambda_j)P_0'(\lambda_j)}[z^n](1-z)^{-2}+o(n).
\end{aligned}$$

Now

$$\begin{aligned}
\sum_{1\le j<d}\frac{2^d}{(2-\lambda_j)P_0'(\lambda_j)} &= \frac{1}{d}\sum_{1\le j<d}\frac{\lambda_j}{2-\lambda_j}\\
&= \frac{2}{d}\sum_{1\le j<d}\frac{1}{2-\lambda_j}-\frac{d-1}{d}\\
&= \frac{P_0''(2)}{dP_0'(2)}-\frac{d-1}{d}\\
&= -\frac{d-1}{2d}.
\end{aligned}$$

Thus

$$\begin{aligned}
&[z^n]\left(\mathbf{I}_{\lambda_{d-1}}\circ\cdots\circ\mathbf{I}_{\lambda_1}\circ\mathbf{I}_2\right)[2^dB](z)\\
&\quad= [z^n]\left(\frac{2}{d}\frac{1}{(1-z)^2}\log\frac{1}{1-z}-\frac{d+3}{2d}\frac{1}{(1-z)^2}\right)+o(n)\\
&\quad= \frac{2}{d}n\log n+\left(\frac{2\gamma}{d}-\frac{1}{2}-\frac{7}{2d}\right)n+o(n),
\end{aligned}\tag{42}$$

since

$$\begin{aligned}
[z^n](1-z)^{-2}\log\frac{1}{1-z} &= (n+1)\sum_{1\le j\le n}j^{-1}-n\\
&= n\log n+(\gamma-1)n+O(\log n).
\end{aligned}$$

The contribution of $2^dC(z)$ and $g(z)$. Similarly, by (27),

$$[z^n]\left(\mathbf{I}_{\lambda_{d-1}}\circ\cdots\circ\mathbf{I}_{\lambda_1}\circ\mathbf{I}_2\right)[2^dC](z) = \frac{2}{d}C^*(2)n+o(n),$$

where $C^*(s) := \int_0^1 C(x)(1-x)^{s-1}\,\mathrm{d}x$, and

$$\begin{aligned}
[z^n]\left(\mathbf{I}_{\lambda_{d-1}}\circ\cdots\circ\mathbf{I}_{\lambda_1}\circ\mathbf{I}_2\right)[g](z) &= \frac{2^{1-d}}{d}[z^n]\mathbf{I}_2[g](z)+o(n)\\
&= \frac{2^{1-d}}{d}g^*(2)n+o(n),
\end{aligned}$$

provided that $g^*(2)$ is finite, where $g^*(s) := \int_0^1(1-x)^{s-1}g(x)\,\mathrm{d}x$.

Boundness of $g^*(2)$. To justify that $g^*(2)$ is finite, we use the same argument as in the proof for $\Lambda(s)$ above. Again by Lemma 2

$$\begin{aligned} g^*(s) &= \sum_{1\le j<d} \int_0^1 (1-x)^{j+s-1} P_j(\vartheta) P_0(\vartheta)^{-1} \left(2^d B(x) + 2^d C(x) + g(x)\right) \, \mathrm{d}x \\ &= \sum_{1\le j<d} \frac{P_j(j+s)}{P_0(j+s)} \int_0^1 (1-x)^{j+s-1} \left(2^d B(x) + 2^d C(x) + g(x)\right) \, \mathrm{d}x \\ &= \sum_{1\le j<d} \frac{P_j(j+s)}{P_0(j+s)} \left(2^d B^*(j+s) + 2^d C^*(j+s) + g^*(j+s)\right), \end{aligned}$$

where

$$B^*(s) = \int_0^1 x(1-x)^{s-3} \, \mathrm{d}x = \frac{1}{(s-1)(s-2)}.$$

Since $B^*(s)$ is finite for $s > 2$, $g^*(s)$ is well-defined for $s > 1$.

Iterating the recurrence as in part (i) gives

$$\begin{aligned} g^*(2) &= \sum_{j\ge 0} V_j \sum_{1\le \ell<d} \frac{P_\ell(j+\ell+2)}{P_0(j+\ell+2)} \left(2^d B^*(j+\ell+2) + 2^d C^*(j+\ell+2)\right) \\ &= \sum_{k\ge 1} \left(2^d B^*(k+2) + 2^d C^*(k+2)\right) \sum_{1\le \ell<d} \frac{P_\ell(k+2)}{P_0(k+2)} V_{k-\ell} \\ &= 2^d \sum_{k\ge 1} \frac{V_k}{k(k+1)} + 2^d \sum_{k\ge 1} V_k C^*(k+2), \end{aligned}$$

where V_k is defined in (17).

Collecting all estimates. Combining this with (42), we obtain

$$A_n = \frac{2}{d} n \log n + K_2 n + o(n),$$

where

$$K_2 = \frac{2\gamma}{d} + \frac{1}{2} - \frac{7}{2d} + \frac{2}{d} \sum_{k\ge 1} \frac{V_k}{k(k+1)} + \frac{2}{d} \sum_{k\ge 0} V_k C^*(k+2).$$

The last series $\sum_{k\ge 0} V_k C^*(k+2)$ is identically zero by the same argument used in part (i) for nonzero initial conditions.

Final simplification. We now show that

$$\sum_{k\ge 1} \frac{V_k}{k(k+1)} = \sum_{1\le j<d} \psi(3-\lambda_j) - (d-1)(1-\gamma), \tag{43}$$

and this will prove (14) by the relations $\psi(3-\lambda_j) = \psi(2-\lambda_j) + (2-\lambda_j)^{-1}$ and

$$\sum_{1\le j<d} \frac{1}{2-\lambda_j} = \frac{d-1}{4}.$$

For that purpose, we substitute the integral representation (38) into the series and then sum over all positive indices k, giving

$$\sum_{k\geq 1}\frac{V_k}{k(k+1)}=\frac{1}{2\pi i}\int_{\varepsilon-i\infty}^{\varepsilon+i\infty}\frac{1}{(s-1)^2}\prod_{1\leq j<d}\frac{\Gamma(3-\lambda_j)\Gamma(s+1)}{\Gamma(s+2-\lambda_j)}\,\mathrm{d}s. \tag{44}$$

Moving the line of integration to the right and taking into account the residue of the unique pole encountered at $s=1$, we obtain (43) by absolute convergence.

A different expression for K_2. Yet another expression for K_2 was derived in [19]

$$K_2=\frac{2\gamma}{d}+\frac{3}{2}-\frac{3}{2d}-2^{d+1}\sum_{k\geq 3}\frac{1}{k(k^d-2^d)}.$$

Equating the two expressions of K_2 leads to the identity

$$2^{d+1}\sum_{k\geq 3}\frac{1}{k(k^d-2^d)}=3-\frac{2}{d}(d-1)\gamma-\frac{2}{d}\sum_{1\leq j<d}\psi(3-\lambda_j)\qquad(d\geq 1),$$

which can be proved using the relations

$$\psi(z+1)=-\gamma+\sum_{j\geq 1}\frac{z}{j(j+z)},$$

(see [14, p.15, Eq. (3)]) and

$$\begin{aligned}\sum_{1\leq j<d}\frac{2-\lambda_j}{k+2-\lambda_j}&=d-1-k\sum_{1\leq j<d}\frac{1}{k+2-\lambda_j}\\&=d-1-k\left(\frac{d(k+2)^{d-1}}{(k+2)^d-2^d}-\frac{1}{k}\right).\end{aligned}$$

Necessity. Consider the case when $A_n=c_0n\log n+c_1n+o(n)$, where $c_0=2/d$. Then, similarly as in part (i), we need the elementary estimate

$$\begin{aligned}[z^n]\mathbf{I}_0[A](z)&=\frac{1}{n}\sum_{0\leq j<n}A_j\\&=\frac{1}{n}\sum_{1\leq j<n}(c_0j\log j+c_1j)+o(n)\\&=\frac{c_0}{2}n\log n+\left(\frac{c_1}{2}-\frac{c_0}{4}\right)n+o(n).\end{aligned}$$

The same estimate holds for $[z^n]z^{-1}\mathbf{I}_0[A](z)$. Iterating the estimates, we obtain

$$[z^n]2^d\left(\mathbf{I}_0\circ\left(z^{-1}\mathbf{I}_0\right)^{d-1}\right)[A](z)=c_0n\log n+\left(c_1-\frac{d}{2}c_0\right)n+o(n).$$

Consequently,

$$B_n=\frac{d}{2}c_0n+o(n)=n+o(n).$$

Thus $B_n-n=o(n)$ and the remaining proof uses the same argument as in part (i). This completes the proof of (13).

2.4 Asymptotic transfer III. Large toll functions

We prove the asymptotic transfer (15) for large toll functions. For general divide-and-conquer recurrences, such a case is always easier than that of small toll functions, one simple reason being that the major contribution comes from a few large terms instead of summing over all small parts like the small toll functions case. More precisely, we expect that most contribution comes from the term $2^d B(z)$ in (22), the other term $g(z)$ being asymptotically negligible.

Assume that $B_n \sim cn^\upsilon$, where $\upsilon > 1$. We start again from (25), which gives

$$\begin{aligned} A_n &= B_n + [z^n]\left(\mathbf{I}_{\lambda_{d-1}} \circ \cdots \mathbf{I}_{\lambda_1} \circ \mathbf{I}_2\right)[g + 2^d B](z) \\ &= B_n + A_n^{[1]} + A_n^{[2]}, \end{aligned}$$

where, by successive applications of (28), we have

$$\begin{aligned} A_n^{[2]} &:= 2^d [z^n]\left(\mathbf{I}_{\lambda_{d-1}} \circ \cdots \mathbf{I}_{\lambda_1} \circ \mathbf{I}_2\right)[B](z) \\ &\sim \frac{c2^d}{P_0(\upsilon+1)} n^\upsilon. \end{aligned}$$

To estimate $A_n^{[1]}$, we first consider $g^*(s) = \int_0^1 (1-x)^{s-1} g(x)\,\mathrm{d}x$, which, by (34), satisfies the recurrence equation

$$\begin{aligned} g^*(s) &= \int_0^1 (1-x)^{s-1}\left(P_0(\vartheta)f(x) - 2^d B(x)\right)\,\mathrm{d}x \\ &= \sum_{1\le j<d} \frac{P_j(j+s)}{P_0(j+s)}\left(g^*(j+s) + 2^d B^*(j+s)\right), \end{aligned} \tag{45}$$

for sufficiently large s. Since $B_n \sim cn^\upsilon$, we deduce that $B^*(s)$ is finite for $s > \upsilon + 1$. The same argument as for $\Lambda(s)$ shows that $g^*(s)$ is finite for $s > \upsilon$. This implies, in particular, that

$$\left|\int_0^1 (1-x)^{\upsilon-\varepsilon} g(x)\,\mathrm{d}x\right| = \left|\Gamma(\upsilon+1-\varepsilon)\sum_{k\ge0} \frac{\Gamma(k+1)}{\Gamma(k+\upsilon+2-\varepsilon)}[z^k]g(z)\right| < \infty.$$

Now by (29) with $\upsilon = 2$

$$[z^n]\mathbf{I}_2[g](z) = (n+1)\sum_{0\le k<n} \frac{[z^k]g(z)}{(k+1)(k+2)}.$$

Let $S_k := \sum_{0\le j\le k} \Gamma(j+1)[z^j]g(z)/\Gamma(j+\upsilon+2-\varepsilon)$. Then $S_k = O(1)$ and, by partial summation,

$$\begin{aligned} (n+1)\sum_{0\le k<n} \frac{[z^k]g(z)}{(k+1)(k+2)} &= (n+1)\sum_{0\le k<n} \frac{\Gamma(k+1)}{\Gamma(k+\upsilon+2-\varepsilon)}[z^k]g(z)\cdot\frac{\Gamma(k+\upsilon+2-\varepsilon)}{\Gamma(k+3)} \\ &= (1-\upsilon+\varepsilon)(n+1)\sum_{0\le k\le n} S_k \frac{\Gamma(k+\upsilon+2-\varepsilon)}{\Gamma(k+4)} + O(n^{\upsilon-\varepsilon}) \\ &= O(n^{\upsilon-\varepsilon}). \end{aligned}$$

Applying now successively (28), we obtain $A_n^{[1]} = O(n^{\upsilon-\varepsilon}) = o(n^\upsilon)$.

From these estimates, it follows that

$$A_n \sim cn^\upsilon + \frac{c2^d}{(\upsilon+1)^d - 2^d} n^\upsilon,$$

which implies the sufficiency part of (15).

Necessity in (15). Assume that $A_n \sim K_3 c n^{\upsilon}$, where $K_3 = (\upsilon+1)^d/((\upsilon+1)^d - 2^d)$. Then, similarly to the necessity proof for case (i),

$$[z^n]2^d \left(\mathbf{I}_0 \circ \left(z^{-1}\mathbf{I}_0\right)^{d-1}\right)[A](z) \sim \frac{2^d}{(\upsilon+1)^d} n^{\upsilon},$$

by successive applications of (28). Then

$$B_n \sim K_3 c\left(1 - \frac{2^d}{(\upsilon+1)^d}\right) n^{\upsilon} \sim c n^{\upsilon}.$$

Simple transfers for the quadtree recurrence (7). The same proof also gives the following O- and o-transfers.

Lemma 4. *Assume $\upsilon > 1$. Then*

$$B_n = O(n^{\upsilon}) \quad \textit{iff} \quad A_n = O(n^{\upsilon}). \tag{46}$$

The same result holds with O replaced by o.

Note that the results for large toll functions can also be proved by other elementary means, but the proof given here based on iterative operators applies for all cases, and is thus more general and uniform.

Recurrence of the Cauchy-Euler part. The preceding analysis shows that when B_n is larger than linear, the contribution from $g(z)$ to A_n is asymptotically negligible. Thus in this case $A_n \sim A_n^{[2]}$, where $P_0(\vartheta)(A^{[2]}(z) - B(z)) = 2^d B(z)$, or in terms of recurrence

$$A_n^{[2]} = B_n + 2^d \sum_{0 \le j < n} \tilde{\pi}_{n,j} A_j^{[2]},$$

where

$$\tilde{\pi}_{n,j} = \frac{1}{n} \sum_{j < j_1 < \cdots < j_{d-1} < n} \frac{1}{j_1 \cdots j_{d-1}},$$

which is to be compared with the alternative expression for $\pi_{n,j}$ (see [19])

$$\pi_{n,j} = \frac{1}{n} \sum_{j < j_1 \le \cdots \le j_{d-1} \le n} \frac{1}{j_1 \cdots j_{d-1}}.$$

2.5 Asymptotic transfer IV. Further refinements

When more precise information on B_n is available, we can refine the preceding approach and obtain more effective approximations to A_n. We consider the following two cases for later use. Recall that $2e^{2\pi i/d} = \alpha + 1 + i\beta$.

Proposition 1. *Assume that A_n satisfies (7).*

(i) *If $B_n \sim c n^{\upsilon}$, where $c, \upsilon \in \mathbb{C}$ and $\alpha < \Re(\upsilon) < 1$, then*

$$A_n = K_B n + \frac{c(\upsilon+1)^d}{(\upsilon+1)^d - 2^d} n^{\upsilon} + o(n^{\Re(\upsilon)} + n^{\varepsilon}),$$

where K_B is defined in (16).

(ii) *If* $B_n = o(n^\alpha)$, *then*

$$A_n = K_B n + K(\lambda_1) n^{\alpha+i\beta} + K(\lambda_2) n^{\alpha-i\beta} + o(n^\alpha + n^\varepsilon), \tag{47}$$

where the $K(\lambda_j)$*'s are defined in (48). If the* B_k*'s are all real, then* $K(\lambda_1) = \overline{K(\lambda_2)}$.

Proof. The proof consists of refining the analysis for the small toll functions part of Theorem 2 using the arguments for large toll functions.

Case (i)**.** Since $B_n \sim cn^\upsilon$, the series in (12) obviously converges. Thus, by (29), we first have

$$\begin{aligned}[z^n]\mathbf{I}_2[g + 2^d B](z) &= (n+1)\left(\sum_{k\ge 0} \frac{g_k + 2^d B_k}{(k+1)(k+2)} - \sum_{k\ge n} \frac{g_k + 2^d B_k}{(k+1)(k+2)}\right) \\ &= K'n - n\sum_{k\ge n} \frac{g_k}{(k+1)(k+2)} - \frac{c2^d}{1-\upsilon} n^\upsilon + o(n^{\Re(\upsilon)}) + O(1),\end{aligned}$$

where $g_k := [z^k]g(z)$ and $K' = \int_0^1 (1-x)\left(g(x) + 2^d B(x)\right) \mathrm{d}x$.

By the same arguments used for $g^*(s)$ in (45), we deduce that $B^*(s)$ is finite for $s > \Re(\upsilon) + 1$ and $g^*(s)$ is bounded for $s > \Re(\upsilon)$. It follows, by the same summation by parts argument used for $A_n^{[1]}$, that

$$n\sum_{k\ge n} \frac{g_k}{(k+1)(k+2)} = O\left(n^{\Re(\upsilon)-\varepsilon}\right).$$

Thus

$$[z^n]\mathbf{I}_2[g + 2^d B](z) = K'n - \frac{c2^d}{1-\upsilon} n^\upsilon + o(n^{\Re(\upsilon)}) + O(1).$$

We may assume that $\Re(\upsilon) > 0$; otherwise all error terms are absorbed in $o(n^\varepsilon)$.

Consider now

$$\begin{aligned}[z^n]\left(\mathbf{I}_{\lambda_1} \circ \mathbf{I}_2\right)[g + 2^d B](z) &= \frac{\Gamma(n+\lambda_1)}{\Gamma(n+1)} \sum_{0\le k<n} \frac{\Gamma(k+1)}{\Gamma(k+1+\lambda_1)}\left(K'k - \frac{c2^d}{1-\upsilon} k^\upsilon + o(k^{\Re(\upsilon)} + k^\varepsilon)\right) \\ &= \frac{K'}{2-\lambda_1} n - \frac{c2^d}{(1-\upsilon)(\upsilon+1-\lambda_1)} n^\upsilon + o(n^{\Re(\upsilon)} + n^\varepsilon),\end{aligned}$$

again by (29). Repeating the same procedure, we obtain

$$A_n - B_n = [z^n]f(z) = \frac{K'}{P_0'(2)} n + \frac{c\,2^d}{P_0(\upsilon+1)} n^\upsilon + o(n^{\Re(\upsilon)} + n^\varepsilon),$$

which proves (i) since $K_B = K'/P_0'(2)$.

Case (ii)**.** Now, similarly as above, we have

$$\begin{aligned}[z^n]\mathbf{I}_2[g + 2^d B](z) &= K'n + o(n^\alpha + n^\varepsilon), \\ [z^n]\mathbf{I}_{\lambda_j}[g + 2^d B](z) &= K_j' n^{\lambda_j - 1} + o(n^\alpha + n^\varepsilon),\end{aligned}$$

where

$$K'_j := \frac{1}{\Gamma(\lambda_j)} \int_0^1 (1-x)^{\lambda_j - 1} \left(g(x) + 2^d B(x)\right) \, \mathrm{d}x \qquad (j = 1, 2).$$

Substituting these estimates into (41) gives

$$\begin{aligned}[z^n] \left(\mathbf{I}_{\lambda_2} \circ \mathbf{I}_{\lambda_1} \circ \mathbf{I}_2\right) [g + 2^d B](z) = \frac{K'}{(2-\lambda_1)(2-\lambda_2)} n + \frac{K'_1}{(\lambda_1 - 2)(\lambda_1 - \lambda_2)} n^{\lambda_1 - 1} \\ + \frac{K'_2}{(\lambda_2 - 2)(\lambda_2 - \lambda_1)} n^{\lambda_2 - 1} + o(n^\alpha + n^\varepsilon).\end{aligned}$$

Applying successively (28) to the remaining operators $\mathbf{I}_{\lambda_j}$ for $j = 3, \ldots d-1$, we obtain (47), where

$$K(\lambda_j) = \frac{2^d}{P'_0(\lambda_j)\Gamma(\lambda_j)} \sum_{k \geq 0} B^*(\lambda_j + k) V_k(\lambda_j) \qquad (j = 1, 2), \tag{48}$$

where $V_k(\lambda_j)$ satisfies the recurrence

$$V_k(\lambda_j) = \sum_{1 \leq \ell < d} \frac{P_i(\lambda_j + \ell)}{P_0(\lambda_j + \ell)} V_{k-\ell}(\lambda_j),$$

with $V_k(\lambda_j) = 0$ if $k < 0$ and $V_0(\lambda_j) = 1$.

The same proof for proving Lemma 3 also implies that $V_k(\lambda_j)$ satisfies the DE

$$\mathbb{D}_z \left(z(1-z)\mathbb{D}_z\right)^{d-1} \left(z^{\lambda_j} V(z)\right) - 2^d z^{\lambda_j - 1} V(z) = 0,$$

and it follows that $V_k(\lambda_j) = O\left(k^{-1} (\log k)^{d-2}\right)$. This justifies the absolute convergence of the series (48). ∎

In a similar way, we also have the following simpler transfer.

Corollary 1. *Assume that $\Re(v) < 1$ and $v \neq \alpha \pm i\beta$. If $B_n = O(n^{\Re(v)})$, then $A_n = K_B n + O(n^{\Re(v)} + n^\alpha + n^\varepsilon)$; if $B_n = o(n^{\Re(v)})$, then $A_n = K_B n + o(n^{\Re(v)}) + O(n^\alpha + n^\varepsilon)$.*

3 Limit laws of X_n: from normal to periodic

We prove first Theorem 1 in this section. Although the first part of Theorem 1 is implied by Theorem 4 below, we give the main steps of the proof by the moment-transfer approach for more logical reasons: first the mean and variance are needed by both proofs (although with different degrees of precision); second, the main hard part of the proof of Theorem 4 consists in refining the estimates of some recursive functionals of moments. We then sketch extensions of the same types of limit results to other toll functions.

The proofs here rely strongly on the different types of asymptotic transfer we developed in Section 2.

3.1 Limit theorems for the number of leaves

Expected number of leaves. By (5), we see that the mean number of leaves in a random quadtree of n nodes satisfies the recurrence (7) with $B_n = \delta_{n,1}$ and $A_0 = 0$. Then $B(x) = x$ and $B^*(s) = s^{-1}(s+1)^{-1}$. Applying (47), we obtain

$$\mathbb{E}(X_n) = \mu_d n + c_+ n^{\alpha + i\beta} + c_- n^{\alpha - i\beta} + o(n^\alpha + n^\varepsilon), \tag{49}$$

for $d \geq 1$, where $c_+ = K(\lambda_1)$ and $c_- = K(\lambda_2)$ with $B^*(s) = s^{-1}(s+1)^{-1}$. In particular,

$$\mu_d = \frac{2}{d} \sum_{k \geq 0} \frac{V_k}{(k+2)(k+3)}.$$

This proves (3) with $G_1(x) = c_+ e^{i\beta x} + c_- e^{-i\beta x}$; see Figure 2 for a plot of the fluctuations of the error terms. We now show that

$$\mu_d = \frac{2^{d+1}}{d} \sum_{k \geq 2} \frac{1}{k^d [k]!} \left((k-1) \sum_{1 \leq j < d} (\psi(k+1-\lambda_j) - \psi(k)) - 2 \right), \tag{50}$$

for $d \geq 2$, which gives an alternative expression to (2).

To prove (50), we apply the integral representation (40), where

$$\begin{aligned} \Upsilon(s) &:= \sum_{k \geq 0} \frac{\Gamma(k+2)\Gamma(1-s)}{(k+2)(k+3)\Gamma(k+2-s)} \\ &= s^2 \psi'(-s) + s - \frac{1}{2} \qquad (\Re(s) < 1). \end{aligned}$$

Now Υ has double poles at all positive integers. Summing over all residues of the double poles of the integrand in (40), we obtain (50) by absolute convergence (since $\Upsilon(s) = O(|s|^{-1})$ as $|s| \to \infty$ and s is at least ε away from all positive integers). Note that

$$(k-1) \sum_{1 \leq j < d} (\psi(k+1-\lambda_j) - \psi(k)) - 2 = d - 1 + O(k^{-1});$$

thus the general terms in (50) decrease at the rate $O(k^{-d})$.

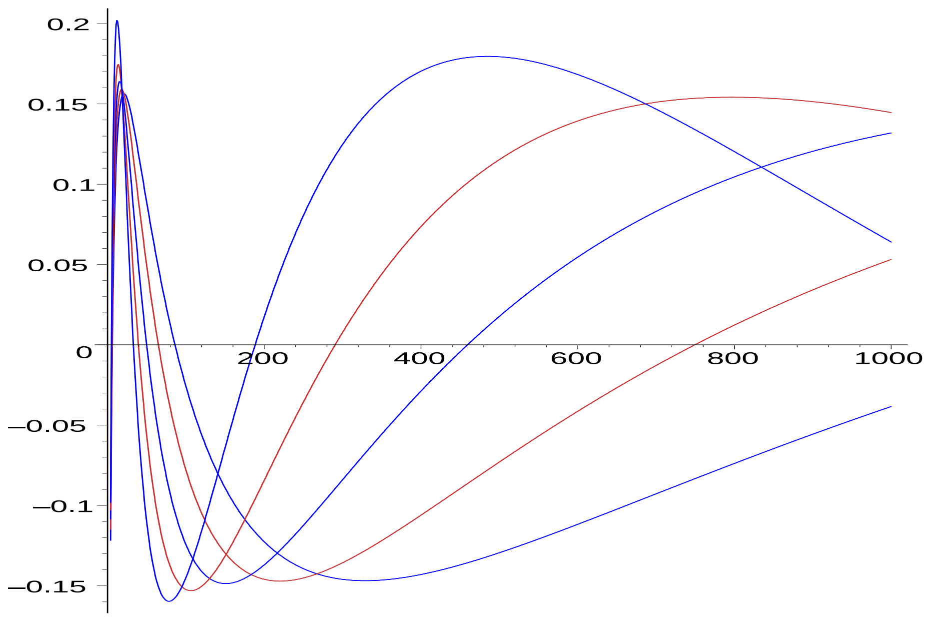

Figure 2: *Periodic fluctuations of* $n^{-\alpha}(\mathbb{E}(X_n) - \mu_d n)$ *for* $n = 4, \ldots, 1000$ *and* $d = 6, \ldots, 10$.

Recurrence of higher moments. For higher moments, we start from the by now standard trick of shifting the mean; thus we consider the moment generating function

$$M_n(y) := \mathbb{E}\left(\exp\left(X_n - \mu_d n - \frac{\mu_d}{2^d-1}\right)y\right),$$

which satisfies, by (5), the recurrence

$$M_n(y) = \sum_{j_1+\cdots+j_{2^d}=n-1} \pi_{n,\mathbf{j}} M_{j_1}(y)\cdots M_{j_{2^d}}(y) \qquad (n \geq 2),$$

with the initial conditions $M_0(y) = e^{-\mu_d y/(2^d-1)}$ and $M_1(y) = e^{(1-2^d\mu_d/(2^d-1))y}$. Note that the additional factor $\mu_d/(2^d-1)$ subtracted has the effect of keeping the recurrence simpler.

Define $M_{n,k} := M_n^{(k)}(0) = \mathbb{E}\left((X_n - \mu_d n - \mu_d/(2^d-1))^k\right)$. Then $M_{n,k}$ satisfies the recurrence

$$M_{n,k} = Q_{n,k} + 2^d \sum_{0\leq j<n} \pi_{n,j} M_{j,k} \qquad (n \geq 2),$$

with the initial conditions $M_{0,k} = (-1)^k \mu_d^k/(2^d-1)^k$ and $M_{1,k} = (1-2^d\mu_d/(2^d-1))^k$, where

$$Q_{n,k} = \sum_{\substack{j_1+\cdots+j_{2^d}=n-1\\ i_1+\cdots+i_{2^d}=k\\ i_1,\dots,i_{2^d}<k}} \binom{k}{i_1,\dots,i_{2^d}} \pi_{n,\mathbf{j}} M_{j_1,i_1}\cdots M_{j_{2^d},i_{2^d}} \qquad (n \geq 2).$$

Note that by (3)

$$M_{n,1} = \begin{cases} O\left(n^\alpha + n^\varepsilon\right), & \text{if } 1 \leq d \leq 8;\\ G_1(\beta\log n)n^\alpha + o(n^\alpha), & \text{if } d \geq 9. \end{cases} \tag{51}$$

Variance. We now prove the asymptotic estimate (4). First we have, by symmetry,

$$Q_{n,2} = 2^{d+1} \sum_{j_1+\cdots+j_{2^d}=n-1} \pi_{n,\mathbf{j}} M_{j_1,1}\left(M_{j_2,1} + \cdots + M_{j_{2^d},1}\right).$$

If $1 \leq d \leq 8$, then the estimate (51) implies that $Q_{n,2} = O(n^{1-2\varepsilon})$. Thus a straightforward application of (12) yields

$$M_{n,2} = \mathbb{E}\left(\left(X_n - \mu_d n - \frac{\mu_d}{2^d-1}\right)^2\right) \sim \sigma_d^2 n,$$

which, by $\mathbb{V}(X_n) = M_{n,2} - M_{n,1}^2$ and (51), implies (4). Here σ_d^2 is given by

$$\sigma_d^2 = \frac{2}{d}\sum_{k,m\geq 0} \frac{V_k m! Q_{m,2}}{(k+2)\cdots(k+m+2)}, \tag{52}$$

with $Q_{0,2}$ and $Q_{1,2}$ properly defined. We will consider numeric evaluations of σ_d^2 later.

If $d \geq 9$, then, by (51),

$$Q_{n,2} = 2^{d+1} \sum_{j_1+\cdots+j_{2^d}=n-1} \pi_{n,\mathbf{j}} \left(c_+ j_1^{\alpha+i\beta} + K(\lambda_2) j_1^{\alpha-i\beta}\right) \times \sum_{2\leq \ell\leq 2^d} \left(c_+ j_\ell^{\alpha+i\beta} + K(\lambda_2) j_\ell^{\alpha-i\beta}\right) + o(n^{2\alpha}).$$

By the strong law of large numbers, we have

$$\begin{aligned}Q_{n,2} = 2^{d+1} \int_{[0,1]^d} \sum_{2\le \ell\le 2^d} \Big(& c_+^2 q_1(\mathbf{x})^{\alpha+i\beta} q_\ell(\mathbf{x})^{\alpha+i\beta} n^{2\alpha+2i\beta} \\ & + c_+c_- \left(q_1(\mathbf{x})^{\alpha+i\beta} q_\ell(\mathbf{x})^{\alpha-i\beta} + q_1(\mathbf{x})^{\alpha-i\beta} q_\ell(\mathbf{x})^{\alpha+i\beta}\right) n^{2\alpha} \\ & + c_-^2 q_1(\mathbf{x})^{\alpha-i\beta} q_\ell(\mathbf{x})^{\alpha-i\beta} n^{2\alpha-2i\beta}\Big) \,\mathrm{d}\mathbf{x} + o(n^{2\alpha}),\end{aligned}$$

where the $q_h(\mathbf{x})$'s are defined in (6). The integrals can be simplified as follows.

$$\begin{aligned}\eta(u,v) &:= \int_{[0,1]^d} q_1(\mathbf{x})^u \sum_{2\le\ell\le 2^d} q_\ell(\mathbf{x})^v \,\mathrm{d}x \\ &= \sum_{0\le \ell<d} \binom{d}{\ell} \left(\frac{1}{u+v+1}\right)^\ell \left(\frac{\Gamma(u+1)\Gamma(v+1)}{\Gamma(u+v+2)}\right)^{d-\ell} \\ &= \left(\frac{1}{u+v+1} + \frac{\Gamma(u+1)\Gamma(v+1)}{\Gamma(u+v+2)}\right)^d - \left(\frac{1}{u+v+1}\right)^d, \end{aligned} \tag{53}$$

for $\Re(u), \Re(v) > -1$. Thus

$$\begin{aligned}Q_{n,2}2^{-d-1} = {} & c_+^2\eta(\alpha+i\beta,\alpha+i\beta)n^{2\alpha+2i\beta} + 2c_-c_+\eta(\alpha+i\beta,\alpha-i\beta)n^{2\alpha} \\ & + c_-^2\eta(\alpha-i\beta,\alpha-i\beta)n^{2\alpha-2i\beta} + o(n^{2\alpha}).\end{aligned}$$

Transferring this approximation term by term using (15) gives

$$M_{n,2} = \tilde{G}_2(\beta\log n)n^{2\alpha} + o(n^{2\alpha}),$$

where

$$\begin{aligned}\tilde{G}_2(u) := {} & 2^{d+1}c_+^2\eta(\alpha+i\beta,\alpha+i\beta)\frac{(2\alpha+2i\beta+1)^d}{P_0(2\alpha+2i\beta+1)}\, e^{2i\beta u} \\ & + 2^{d+2}c_-c_+\eta(\alpha+i\beta,\alpha-i\beta)\frac{(2\alpha+1)^d}{P_0(2\alpha+1)} \\ & + 2^{d+1}c_-^2\eta(\alpha-i\beta,\alpha-i\beta)\frac{(2\alpha-2i\beta+1)^d}{P_0(2\alpha-2i\beta+1)}\, e^{-2i\beta u}.\end{aligned}$$

This proves (4) with $G_2(x) = \tilde{G}_2(x) - G_1(x)^2$.

Asymptotic normality for $1 \le d \le 8$. The same arguments used above for the variance also apply for $M_{n,k}$ for $k \ge 3$. By induction, we obtain

$$\begin{cases} M_{n,2k} \sim \dfrac{(2k)!}{2^k k!}\,\sigma_d^{2k} n^{2k}; \\ M_{n,2k-1} = o(n^{k-1/2}), \end{cases}$$

for $k \ge 1$; details are omitted here for conciseness; see [3] for a similar proof. This proves the first part of Theorem 1.

Periodic fluctuations for $d \geq 9$. In this case, the same calculations for $\mathbb{V}(X_n)$ can be extended to show that

$$\mathbb{E}\left(\left(X_n - \mu_d n - \frac{\mu_d}{2^d - 1}\right)^k\right) \sim \tilde{G}_k(\beta \log n) n^{k\alpha} \qquad (k \geq 2); \tag{54}$$

where the $\tilde{G}_k$'s are bounded periodic functions. Then the proof that there is no fixed limit law for $(X_n - \mathbb{E}(X_n))/\sqrt{\mathbb{V}(X_n)}$ follows the same arguments used in [3].

Instead of giving the messy details of the proof for (54), we sketch the proof for

$$\|X_n - \mu_d n - 2\Re(n^{\alpha+i\beta} X)\|_p = o(n^\alpha) \qquad (p \geq 2), \tag{55}$$

where $\|Z\| = (\mathbb{E}|X|^p)^{1/p}$ denotes the usual L_p norm. Here X is a random variable with $\mathbb{E}(X) = c_+$ (see (49)) and defined by

$$X \stackrel{\mathscr{D}}{=} \langle U \rangle_1^{\alpha+i\beta} X^{(1)} + \cdots + \langle U \rangle_{2^d}^{\alpha+i\beta} X^{(2^d)},$$

where the $X^{(i)}$'s are independent copies of X and the $\langle U \rangle_i$'s are the volumes of the 2^d quadrants split by a random point in $[0,1]^d$. Part (ii) of Theorem 1 also follows from (55).

It suffices to prove $p = 2$, the remaining cases following by induction. The arguments used here are modified from those in [15] for random m-ary search trees.

Define

$$\begin{aligned}
\xi_n &:= \left\| X_n - \mu_d n - 2 \sum_{1 \leq j \leq 2^d} \Re\left(J_j^{\alpha+i\beta} X^{(j)}\right)\right\|_2, \\
\eta_n &:= \left\| 2 \sum_{1 \leq j \leq 2^d} \Re\left(J_j^{\alpha+i\beta} X^{(j)}\right) - 2 \sum_{1 \leq j \leq 2^d} \Re\left(n^{\alpha+i\beta} \langle U \rangle_j^{\alpha+i\beta} X^{(j)}\right)\right\|_2.
\end{aligned}$$

We prove that $\xi_n, \eta_n = o(n^\alpha)$, which will then imply (55) for $p = 2$.

First by the decomposition

$$\xi_n \leq \|X_n - \mu_d n\|_2 + 2^{d+2} \|J_1^{\alpha+i\beta} X^{(1)}\|_2,$$

we deduce that $\xi_n = O(n^\alpha)$. Then by the recurrence (5), we have the inequality

$$\xi_n^2 \leq \sum_{1 \leq j \leq 2^d} \mathbb{E}\left(\xi_{J_j} + \eta_{J_j}\right)^2 + o(n^{2\alpha}).$$

This, together with the estimate

$$\eta_n \leq 2^{d+2} n^\alpha \|X^{(1)}\|_2 \left\| \left(\frac{J_1}{n}\right)^{\alpha+i\beta} - \langle U \rangle_2^{\alpha+i\beta} \right\|_2 = o(n^\alpha),$$

gives

$$\begin{aligned}
\xi_n^2 &\leq 2^d \sum_{0 \leq j < n} \pi_{n,j} \xi_j^2 + o(n^{2\alpha}) \\
&= o(n^{2\alpha}),
\end{aligned}$$

by the o-version of (46).

d	$\mu_d \approx$
2	0.47841 76043 57434 47533 79639 99504 60454 12547 97628
3	0.56850 70194 06572 68270 35257 03246 03680 11920 50021
4	0.63168 48783 52998 69050 68769 97892 90145 67365 77851
5	0.67906 23676 94926 62299 74554 08602 48628 92348 92646
6	0.71615 83294 69847 70674 65510 61878 16738 93088 58805
7	0.74609 46112 09331 64803 70711 94105 57503 99390 36451
8	0.77079 60778 85838 99509 15248 99261 83895 90393 54520
9	0.79152 59978 40106 48407 81034 62942 59540 22737 03660
10	0.80915 45900 27608 17078 62137 34456 57737 58997 15908

Table 2: Approximate numeric values of μ_d for $d = 2, \dots, 10$.

3.2 Numerics of μ_d and σ_d^2

We consider means of computing numerically the constants μ_d and σ_d^2.

Numerical values of μ_d. To compute the constants μ_d to high precision, one can use either (2) or (50) by the standard procedure: compute the first few terms exactly and estimate the remaining terms by their asymptotic behaviors.

An alternative procedure is described in the last section. Consider $\bar{f}(z) := f(z) - \sum_{2 \le j < N} A_j z^j$ ($A_1 = B_1$ and $B_n = 0$ for $n \ge 2$) for a suitably large number N, say 50. Exact values of A_n can be easily computed by the exact expression (1) when n is small. Observe that

$$\vartheta(z\vartheta)^{d-1} \sum_{j \ge N} c_j z^j = \sum_{j \ge N-1} c'_j z^j.$$

Thus the right-hand side of the DE

$$\left(\vartheta(z\vartheta)^{d-1} - 2^d\right) \bar{f} = 2^d z - \left(\vartheta(z\vartheta)^{d-1} - 2^d\right) \sum_{2 \le j < N} A_j z^j,$$

contains only monomials z^j with $N < j < N + d$. Then the new $B^*(s)$ is of order s^{-N} for large s, implying a better convergence rate for the series (16) since V_k remains the same and can be computed recursively. Then we need only compute the first few terms (10 for example) of the series (16) to give the required degree of precision. In this way, we obtain Table 3.2. Such a procedure is also useful for other constants such as σ_d^2.

Expressions for σ_d^2. We first derive more explicit expressions for $M_{n,2}$ in (52) before computing σ_d^2.

We start from the bivariate generating function $F(z, y) := \sum_{n \ge 0} \mathbb{E}(e^{X_n y}) z^n / n!$, which satisfies, by (5), the equation

$$\frac{\partial}{\partial z} F(z, y) = e^y - 1 + \int_{[0,1]^d} F(q_1(\mathbf{x})z, y) \cdots F(q_{2^d}(\mathbf{x})z, y) \, \mathrm{d}\mathbf{x}.$$

In particular, $F(z, 0) = e^z$.

Then the Poisson generating function

$$\tilde{F}(z,y) = e^{-z}\sum_{n\geq 0} M_n(y)\frac{z^n}{n!} = e^{-z}\sum_{n\geq 0}\mathbb{E}(e^{(X_n-\mu_d n-\mu_d/(2^d-1))y})\frac{z^n}{n!}$$

satisfies the equation

$$\tilde{F}(z,y) + \frac{\partial}{\partial z}\tilde{F}(z,y) = e^{-z}(e^y-1)e^{-2^d\mu_d y/(2^d-1)} + \int_{[0,1]^d}\tilde{F}(q_1(\mathbf{x})z,y)\cdots\tilde{F}(q_{2^d}(\mathbf{x})z,y)\,\mathrm{d}\mathbf{x}.$$

Let $\tilde{F}(z,y) = \sum_{j\geq 0}\tilde{F}_j(z)y^j/j!$. Then

$$\tilde{F}_1'(z) + \tilde{F}_1(z) = e^{-z} + 2^d\int_{[0,1]^d}\tilde{F}_1(x_1\cdots x_d z)\,\mathrm{d}\mathbf{x},$$

with the initial condition $\tilde{F}_1(0) = -\mu_d/(2^d-1)$. The coefficients $u_n := n![z^n]\tilde{F}_1(z)$ satisfy

$$u_{n+1} + u_n = (-1)^n + \frac{2^d}{(n+1)^d}\,u_n,$$

which, after iterating, can be solved to be

$$u_n = (-1)^{n-1}\sum_{2\leq k\leq n}\prod_{k<\ell\leq n}\left(1-\frac{2^d}{\ell^d}\right) = (-1)^{n-1}[n]!\sum_{2\leq k\leq n}\frac{1}{[j]!},$$

for $n\geq 2$, with $u_0 = -\mu_d/(2^d-1)$ and $u_1 = 1-\mu_d$.

For $\tilde{F}_2(z)$, we have the same type of equation

$$\tilde{F}_2'(z) + \tilde{F}_2(z) = \tilde{g}_2(z) + 2^d\int_{[0,1]^d}\tilde{F}_2(x_1\cdots x_d z)\,\mathrm{d}\mathbf{x},$$

with the initial condition $\tilde{F}_2(0) = \mu_d^2/(2^d-1)^2$, where

$$\tilde{g}_2(z) := \left(1-\frac{2^{d+1}\mu_d}{2^d-1}\right)e^{-z} + 2^d\int_{[0,1]^d}\tilde{F}_1(x_1\cdots x_d z)\sum_{2\leq\ell\leq 2^d}\tilde{F}_1(q_\ell(\mathbf{x})z)\,\mathrm{d}\mathbf{x}. \tag{56}$$

Observe that

$$n![z^n]2^d\int_{[0,1]^d}\tilde{F}_1(x_1\cdots x_d z)\sum_{2\leq\ell\leq 2^d}\tilde{F}_1(q_\ell(\mathbf{x})z)\,\mathrm{d}\mathbf{x} = 2^d\sum_{0\leq j\leq n}\binom{n}{j}u_j u_{n-j}\eta(j,n-j),$$

where $\eta(j,n-j)$ is defined in (53).

By (56), we then have for $n\geq 0$

$$v_n := n!(-1)^n[z^n]\tilde{g}_2(z) = 1-\frac{2^{d+1}\mu_d}{2^d-1} + 2^d(-1)^n\sum_{0\leq j\leq n}\binom{n}{j}u_j u_{n-j}\eta(j,n-j).$$

It follows that

$$n![z^n]\tilde{F}_2(z) = (-1)^{n-1}[n]!\sum_{1\leq k<n}\frac{v_k}{[k+1]!}\qquad (n\geq 2),$$

d	$\sigma_d^2 \approx$
2	0.06145 73978 66984 07284 36701 54743 66750 63784
3	0.06802 65800 83909 72781 61723 15284 91262 75906
4	0.07090 19719 94546 02309 70950 30497 53882 55032
5	0.07261 12472 86535 68765 26637 38060 39503 98071
6	0.07449 21253 93111 00674 61761 51696 97039 29930
7	0.07731 76983 93655 71830 91768 87307 89088 95507
8	0.08123 98836 52827 96294 47650 19430 64044 32562

Table 3: Approximate numeric values of σ_d^2 for $d = 2, \ldots, 8$. Note that $\sigma_1^2 = 2/45 \approx 0.04444\ldots$.

with $\tilde{F}_2(0) = \mu_d^2/(2^d-1)^2$ and $\tilde{F}_2'(0) = 1 - 2^{d+1}\mu_d/(2^d-1) + (2^d+1)\mu_d^2/(2^d-1)$, and consequently

$$\begin{aligned} M_{n,2} &= \mathbb{E}\left(X_n - \mu_d\left(n + \frac{1}{2^d-1}\right)\right)^2 \\ &= \frac{\mu_d^2}{(2^d-1)^2} + \left(1 - \frac{2^{d+1}\mu_d}{2^d-1} + \frac{2^d+1}{2^d-1}\mu_d^2\right) n - \sum_{2\le k\le n} \binom{n}{k}(-1)^k [k]! \sum_{1\le j<k} \frac{v_j}{[j+1]!}. \end{aligned}$$

This provides a less dimension dependent expression for computing $M_{n,2}$ for small values of n needed for computing the approximate values of σ_d^2 in Table 3.2.

Note that for $1 \le d \le 8$, $M_{n,1} = O(n^{0.42})$ and

$$\mathbb{V}(X_n) = M_{n,2} - M_{n,1}^2 = \mathbb{E}\left(X_n - \mu_d\left(n + \frac{1}{2^d-1}\right)\right)^2 - M_{n,1}^2;$$

Thus to compute the limiting constant σ_d^2 of $\mathbb{V}(X_n)/n$, it suffices to compute $M_{n,2}$.

By the same procedure for computing μ_d, we obtain Table 3.2.

Note that

$$Q_{n,2} = [z^{n-1}]e^z\tilde{g}_2(z) = \sum_{0\le j<n} \binom{n-1}{j}(-1)^j v_j \qquad (n \ge 1).$$

For consistency, we can define $Q_{0,2} := \mu_d^2/(2^d-1)^2$. Then $Q_{1,2} = v_0 = 1 - 2^{d+1}\mu_d/(2^d-1) + 2^d\mu_d^2/(2^d-1)$ and for $n \ge 2$

$$Q_{n,2} = 2^d \sum_{0\le m<n} \binom{n-1}{m} \sum_{0\le j\le m} \binom{m}{j} u_j u_{m-j} \eta(j, m-j).$$

3.3 Phase change of other cost measures

Consider the random variables defined recursively by

$$Y_n \overset{\mathscr{D}}{=} Y_{J_1}^{(1)} + \cdots + Y_{J_{2^d}}^{(2^d)} + T_n \qquad (n \ge 1), \tag{57}$$

with Y_0 given, where the $(Y_n^{(i)})$'s are independent copies of Y_n and T_n is a known random variable (often called "toll function").

3.3.1 Phase change of general toll functions

Our method of proof extends easily to cover a wide class of toll functions. We formulate a simple result for deterministic toll functions as follows.

Theorem 3. *If $T_n = O(n^{1/2}(\log n)^{-1/2-\varepsilon})$ and T_n is not identically 1 for all $n \geq 1$, then*

$$\frac{Y_n - \mu'_d n}{\sigma'_d \sqrt{n}} \xrightarrow{\mathscr{D}} N(0,1),$$

for $1 \leq d \leq 8$, where μ'_d and σ'_d are constants; if $d \geq 9$, then the sequence of random variables $(Y_n - \mathbb{E}(Y_n))/\sqrt{\mathbb{V}(Y_n)}$ does not converge to a fixed limit law.

The proof follows from that for Theorem 1 and is omitted. Both constants μ'_d and σ'_d can be computed by the same procedure as for μ_d and σ_d.

By the recurrence

$$\mathbb{V}(Y_n) = \sum_{0 \leq j < n} \pi_{n,j} \left(\mathbb{E}(Y_{j_1}) + \cdots + \mathbb{E}(Y_{j_{2^d}}) - \mathbb{E}(Y_n) + T_n\right)^2 + 2^d \sum_{0 \leq j < n} \pi_{n,j} \mathbb{V}(Y_j),$$

we see that the variance is identically zero iff $T_n \equiv 1$ for $n \geq 1$. In this case, $Y_n \equiv n$ (the total number of nodes in the tree). This also implies, when applying (12), the identity

$$\frac{2}{d} \sum_{k \geq 0} \frac{V_k}{(k+1)(k+2)} = 1 \qquad (d \geq 1). \tag{58}$$

The same method of proof we used for proving Theorem 1 also applies to cover the case when $T_n \sim \sqrt{n}$, which still leads to asymptotic normality for Y_n when $1 \leq d \leq 8$ with linear mean but with variance of order $n \log n$. The same non-existence of fixed limit law also holds in the wider range $T_n = o(n^\alpha)$ when $d \geq 9$. More cases can be clarified as in [7]. Since the number of concrete examples (directly related to cost measures of algorithms or quadtrees) is limited, we stop from considering other general limit results.

3.3.2 Concrete examples and extensions

We briefly discuss instead a few instances of T_n studied before in the literature.

Paging. The page usage of random quadtrees was studied in [26] and [19]; it can be regarded as a generalization of the number of leaves and satisfies (57) with $T_n = 1$ when $n > b$, and $T_n = 0$ otherwise, where $b \geq 0$ is a predetermined structural constant. We can also view Y_n as enumerating the number of nodes x with subtree sizes rooted at x larger than b.

By Theorem 3, *the page usage in random quadtrees undergoes the same type of phase change (of limit laws) as the number of leaves*. The mean constant is given by

$$\mu'_d(b) = \frac{2}{d} \sum_{k \geq 0} \frac{(b+1)! V_k}{(k+1)(k+2) \cdots (k+b+2)}.$$

If $d = 2$, then (see (35))

$$\begin{aligned}\mu_2'(b) &= 12(b+1)! \sum_{k\geq 0} \frac{(k+1)!}{(k+3)(k+4)(k+b+2)!} \\ &= 12(b+1) \int_0^1 (1-x)^b x^{-3} \left((1-x)\log(1-x) + x - \frac{x^2}{2} - \frac{x^3}{6} \right) \mathrm{d}x \\ &= 6b^2 + 9b + 1 - b(b+1)^2\pi^2 + 6b(b+1)^2 \sum_{1\leq j\leq b} j^{-2},\end{aligned}$$

which coincides with the expression first derived in [26].

For $d \geq 3$, expressions for μ_d' are less explicit. We first simplify $\Upsilon(s)$ (see (40)) as follows.

$$\begin{aligned}\Upsilon(s) &= \sum_{k\geq 0} \frac{(b+1)!}{(k+2)\cdots(k+b+2)} \cdot \frac{\Gamma(k+1)\Gamma(1-s)}{\Gamma(k+2-s)} \\ &= (b+1) \sum_{0\leq \ell\leq b} \binom{b}{\ell} (-1)^\ell \Omega_{\ell+2}(s),\end{aligned}$$

where

$$\Omega_a(s) := \int_0^1 (1-x)^{-s} \sum_{k\geq 0} \frac{x^k}{k+a} \,\mathrm{d}x \qquad (\Re(s) < 1;\ a = 0, 1, \dots),$$

(when $a = 0$, the term corresponding to $k = 0$ is dropped). Obviously, $\Omega_0(s) = (s-1)^{-2}$, and

$$\Omega_1(s) = \sum_{k\geq 1} \frac{1}{(s-k)^2} = \psi'(1-s) \qquad (\Re(s) < 1).$$

By an integration by parts, we have the recurrence

$$\Omega_{a+1}(s) = \frac{s}{a}\,\Omega_a(s+1) + \frac{1}{a^2} - \frac{1}{as} \qquad (a \geq 1).$$

By induction

$$\Omega_a(s) = \binom{s+a-2}{a-1} \psi'(1-s) + \mathrm{poly}_1(a; s) \qquad (a = 1, 2, \dots),$$

where $\mathrm{poly}_1(a; s)$ is a polynomial of degree $a-2$ such that $\Omega_a(s)$ is of growth order $|s|^{-1}$ at infinity (with $|s-k| \geq \varepsilon$). More precisely, since

$$\psi'(1-s) = \sum_{j\geq 0} (-1)^{j+1} \mathbf{B}_j s^{-j-1} \qquad (|s| \to \infty, |\arg(-s)| \leq \pi - \varepsilon),$$

where the $\mathbf{B}_j$'s denote Bernoulli numbers (see [14, p. 47, Eq. (7)])

$$\mathrm{poly}_1(a; s) = \sum_{1\leq j<a} s^{j-1} \sum_{j\leq \ell<a} \frac{|\mathbf{s}(a-1,\ell)|}{(a-1)!} (-1)^{j-\ell} \mathbf{B}_{j-\ell} \qquad (a \geq 2),$$

where the $\mathbf{s}(a-1, j)$'s denote Stirling numbers of the first kind. From this expression, we deduce the representation

$$\Upsilon(s) = \frac{(-1)^b}{b!} s(s-1)\cdots(s-b)\psi'(1-s) + \mathrm{poly}_2(b; s),$$

where $\mathrm{poly}_2(b;s)$ is a polynomial of degree b such that $\Upsilon(s)$ is of growth order $|s|^{-1}$ at infinity (with $|s-k| \geq \varepsilon$).

Then the integrand in the integral (40) has simple poles at $s = 1, 2, \dots, b$ and double poles at $s = b+1, b+2, \dots$. Summing over all residues of the poles yields

$$\begin{aligned}\mu_d'(b) &= \frac{2^{d+1}}{d} \sum_{1 \leq k \leq b} \frac{(-1)^k}{\binom{b}{k}(k+1)^d k[k+1]!} \\ &\quad + \frac{2^{d+1}}{d} \sum_{k>b} \frac{(-1)^b (b+1)\binom{k}{b+1}}{(k+1)^d k[k+1]!} \left(\sum_{1 \leq j < d} (\psi(k+2-\lambda_j) - \psi(k+1)) - \psi(k+1) + \psi(k-b) \right).\end{aligned}$$

Note that the last series diverges for $b \geq d$. Numerically, the procedure we used for computing μ_d is preferable.

When $b \geq d$, we can use the recurrence

$$\mu_d'(b) = 2^{-d} \sum_{0 \leq j \leq d} R_{d,j} \mu_d'(b+j-1) \qquad (b \geq 1), \tag{59}$$

so that once the values $\{\mu_d'(0), \dots, \mu_d'(d-1)\}$ are known, all values of $\mu_d'(b)$ for higher values of b can be computed successively. Here $R_{d,j}$ is defined recursively as $R_{0,0} := 1$ and

$$R_{d,j} = (b+j+1) R_{d-1,j} - (b+j-1) R_{d-1,j-1} \qquad (0 \leq j \leq d), \tag{60}$$

with $R_{d,j} = 0$ when $j < 0$ or $j > d$. The recurrence (59) is proved using the DE (37) and successive integration by parts as follows.

$$\begin{aligned}\mu_d'(b) &= \frac{2}{d} \int_0^1 (1-x)^{b+1} V(x)\,\mathrm{d}x \\ &= \frac{2^{1-d}}{d} \int_0^1 \frac{(1-x)^b}{x^2} \left(x(1-x)\mathbb{D}\right)^d x^2 V(x)\,\mathrm{d}x \\ &= \frac{2^{1-d}}{d} \int_0^1 (1-x)^b R_d(x) V(x)\,\mathrm{d}x,\end{aligned}$$

where $R_d(x) = R_d(b;x)$ is defined by

$$\begin{aligned}R_d(x) &:= \frac{x^2}{(1-x)^b} \left(-\mathbb{D}x(1-x)\right)^d \frac{(1-x)^b}{x^2} \\ &= \sum_{0 \leq j \leq d} R_{d,j} (1-x)^j,\end{aligned}$$

with $R_{d,j}$ satisfying (by induction) the recurrence (60). Thus (59) follows. Note that when $b = 0$

$$\mu_d'(0) = \frac{2}{d} \int_0^1 (1-x) V(x)\,\mathrm{d}x = \frac{2^{1-d}}{d} \int_0^1 V(x)\,\mathrm{d}x = 1,$$

which can be proved directly by (40); see also (58).

Node sorts. If T_n is equal to the probability that the root has b nonempty subtrees, where $0 \le b \le 2^d$, then Y_n represents the number of nodes in random quadtrees having exactly b nonempty subtrees. The same type of phase change phenomenon holds since the toll function is bounded; see [34, 35] for expressions for the probability the root having b subtrees.

In general, if $T_n = \delta_{n,b}$, where $b \ge 0$, then the limits $\mu_d' = \mu_d'(b)$ of $\mathbb{E}(Y_n)/n$ are called *universal constants* in [36] since for general toll functions T_n with linear mean the linearity constant can be expressed in terms of the $\mu_d'(b)$'s as $\sum_{b\ge 1} T_b \mu_d'(b)$. Expressions for $\mu_d'(b)$ can be derived similar to the previous case. We have

$$\begin{aligned}\Upsilon(s) = \Upsilon_b(s) &= \sum_{k\ge 0} \frac{b!\Gamma(k+1)\Gamma(1-s)}{(k+2)\cdots(k+b+2)\Gamma(k+2-s)} \\ &= -\sum_{0\le \ell\le b} \binom{b}{\ell}(-1)^\ell(\ell+1)\Omega_{\ell+2}(s) \\ &= (-1)^{b+1}\frac{s^2(s-1)\cdots(s-b+1)}{b!}\psi'(1-s) + \mathrm{poly}_3(b;s),\end{aligned}$$

where $\mathrm{poly}_3(b;s)$ is a polynomial of degree b such that $\Upsilon(s)$ is of growth order $|s|^{-1}$ at infinity (with $|s-k| \ge \varepsilon$). Also $\mu_d'(b)$ satisfies the recurrence

$$\mu_d'(b) = 2^{-d}\sum_{0\le j\le d} R_{d,j}\mu_d'(b+j-1) \qquad (b \ge 1),$$

with $R_{d,j}$ satisfying $R_{d,j} = (b+j)R_{d-1,j} - (b+j-1)R_{d-1,j-1}$ for $0 \le j \le d$. Note that in this case $R_{d,0} = b^d$ and $R_{d,j} = (-1)^{d-1}(P_{j-1}(-b) - P_j(-b))$ for $1 \le j \le d$.

Total path length. In this case, $T_n = n-1$. Although Theorem 3 does not apply, our method of moments does, and we obtain convergence of all moments of $(Y_n - \mathbb{E}(Y_n))/n$ to some non-normal limit law for each $d \ge 1$; see [40], and [30] for similar details. In particular, the mean satisfies (see (14))

$$\mathbb{E}(Y_n) \sim \frac{2}{d}n\log n - \left(2 + \frac{2}{d} - 2\gamma - \frac{2}{d}\sum_{1\le j<d}\psi(2-\lambda_j)\right)n,$$

and the variance is asymptotic to $K_4 n^2$, where

$$K_4 = \frac{3^d}{3^d - 2^d}\int_{[0,1]^d}\left(1 + \frac{2}{d}\sum_{1\le j\le 2^d} q_j(\mathbf{x})\log q_j(\mathbf{x})\right)^2 \mathrm{d}\mathbf{x}.$$

To evaluate the integral, let

$$\tilde{\eta}(u,v) = \int_{[0,1]^d} q_1(\mathbf{x})^u \sum_{1\le \ell\le 2^d} q_\ell(\mathbf{x})^v \,\mathrm{d}\mathbf{x}.$$

Then $\tilde{\eta}(u,v) = \eta(u,v) + 1/(u+v+1)^d$, where η is defined in (53), so that

$$\tilde{\eta}(u,v) = \left(\frac{1}{u+v+1} + \frac{\Gamma(u+1)\Gamma(v+1)}{\Gamma(u+v+2)}\right)^d.$$

It follows that

$$
\begin{aligned}
K_4 &= \frac{3^d}{3^d-2^d}\left(1+\frac{4}{d}\cdot\frac{\partial}{\partial v}\,\tilde{\eta}(0,v)\Big|_{v=1}+\frac{4}{d^2}\,2^d\,\frac{\partial^2}{\partial u\partial v}\,\tilde{\eta}(u,v)\Big|_{u=1,v=1}\right)\\
&= \frac{3^d}{3^d-2^d}\cdot\frac{21-2\pi^2}{9d};
\end{aligned}
$$

see also [40].

Unlike the number of leaves and other small cost measures, there is no change of limit law for total path length since the order of the variance is not alterned for increasing d.

Expected profiles (or depth). Denote by $Z_{n,k}$ the number of nodes at distance k to the root; the $Z_{n,k}$'s are informative shape characteristics often referred to as the *profiles* of the trees. The *depth* D_n is the distance of a randomly chosen node (all n nodes being equally likely) to the root. Then the probability that the depth is k equals $\mathbb{E}(Z_{n,k})/n$. Consider the level polynomials $L_n(y) := \sum_k \mathbb{E}(Z_{n,k})y^k$. Then $L_n(y)$ satisfies the recurrence

$$
L_n(y) = 1 + 2^d y \sum_{0\le j<n} \pi_{n,j} L_j(y) \qquad (n\ge 1),
$$

with $L_0(y)=0$; see [19]. The same analysis for the small toll functions part of Theorem 2 (and the error analysis in Section 2.5) applies *mutatis mutandis* and yields

$$
L_n(y) = \mathcal{K}(y) n^{2y^{1/d}-1} + O\left(n^{2\Re(y^{1/d}e^{2\pi i/d})-1} + n^{\varepsilon}\right), \tag{61}
$$

where the O-term holds uniformly for y lying in some complex neighborhood of unity, and

$$
\mathcal{K}(y) = \frac{2^d y^{1/d}}{d}\sum_{k\ge 2}\frac{\prod_{3\le \ell\le k}(1-2y^{1/d}/\ell)}{k^{d-1}\prod_{3\le\ell\le k}(1-2^d y/\ell^d)}\left((k-1)\sum_{1\le j<d}\left(\psi(k+1-\lambda_j y^{1/d})-\psi(k)\right)-1\right).
$$

Thus the asymptotic normality (with optimal Berry-Esseen bound) of the depth D_n follows from (61) and the so-called quasi-power approximation theorems; see [24, Sec. IX.5] or [27]. Note that

$$
\mathcal{K}(1) = \frac{2^{d+1}}{d}\sum_{k\ge 2}\frac{1}{k^d[k]!}\left(\sum_{1\le j<d}\left(\psi(k+1-\lambda_j)-\psi(k)\right)-\frac{1}{k-1}\right) = 1 \qquad (d\ge 2);
$$

compare (58).

A considerable simplification of the expression for $\mathcal{K}(y)$ can be obtained by applying the finite difference integral representation for the closed-form expression (see [19])

$$
L_n(y) = n - (1-y)\sum_{2\le k\le n}\binom{n}{k}(-1)^k\prod_{3\le j\le k}\left(1-\frac{2^d}{j^d}\,y\right) \qquad (n\ge 0),
$$

giving

$$
L_n(y) = -\frac{1}{2\pi i}\int_{\frac12-i\infty}^{\frac12+i\infty}\frac{\Gamma(n+1)\Gamma(-s)}{\Gamma(n+1-s)\Gamma(s+1)^d}\prod_{0\le\ell<d}\frac{\Gamma(s+1-\lambda_\ell y^{1/d})}{\Gamma(2-\lambda_\ell y^{1/d})}\,\mathrm{d}s.
$$

Then, by moving the line of integration to the left and summing the simple poles encountered, we obtain

$$
L_n(y) = \frac{1}{1-2^d y} + \mathcal{K}(y) n^{2y^{1/d}-1}\left(1+O\left(n^{-\varepsilon}+n^{-2\Re(y^{1/d}(1-e^{2\pi i/d}))}\right)\right),
$$

uniformly for $|y| \geq 2^{-d} + \varepsilon$, where

$$\mathcal{K}(y^d) := \frac{1}{\Gamma(2y)^d(2y-1)} \prod_{1\leq \ell<d} \frac{\Gamma(2y(1-e^{2\ell\pi i/d}))}{\Gamma(2-2ye^{2\ell\pi i/d})}.$$

This explicit expression and the quasi-power theorems in [27] also give more precise estimates for the mean and variance of the depth

$$\begin{aligned}
\mathbb{E}(D_n) &= \frac{2}{d}\log n + [t]\log \mathcal{K}(e^t) + o(1),\\
\mathbb{V}(D_n) &= \frac{2}{d^2}\log n + 2[t^2]\log \mathcal{K}(e^t) + o(1),
\end{aligned}$$

where

$$\begin{aligned}
[t]\log \mathcal{K}(e^t) &= K_2 - 1 = -2 - \frac{2}{d} + 2\gamma + \frac{2}{d}\sum_{1\leq j<d} \psi(2-\lambda_j),\\
2[t^2]\log \mathcal{K}(e^t) &= \frac{2}{d}(1+\gamma) - \frac{2\pi^2}{3d} + \frac{2}{d^2} + \frac{2}{d^2}\sum_{1\leq j<d} \left(\psi(2-\lambda_j) + 2(1-\lambda_j)\psi'(2-\lambda_j)\right).
\end{aligned}$$

Note that $n\mathbb{E}(D_n)$ equals the expected total path length, or A_n when $B_n = n-1$.

4 Second phase change: convergence rates and local limit theorems for X_n

We consider the convergence rate and local limit theorem for X_n, which undergo another phase change. Local limit theorems are more informative and precise than asymptotic normality. We use characteristic functions and standard Fourier analysis (see [42]), the main estimate needed being based on the refined method of moments introduced in [28] and the refined asymptotic transfers developed in Section 2.5.

Local limit theorems. To state our result, let

$$\bar{\alpha} := \begin{cases} 1/3, & \text{if } 1 \leq d \leq 7;\\ \sqrt{2}-1, & \text{if } d = 8. \end{cases}$$

Theorem 4. *Uniformly for $x = o(n^{1/2-\bar{\alpha}})$,*

$$\mathbb{P}\left(X_n = \left\lfloor X_n + x\sqrt{\mathbb{V}(X_n)}\right\rfloor\right) = \frac{e^{-x^2/2}}{\sqrt{2\pi\mathbb{V}(X_n)}}\left(1 + O\left((1+|x|^3)n^{-3(1/2-\bar{\alpha})}\right)\right).$$

The error terms in both cases are, up to the implied constants, optimal. Numerically, $3(1/2-\bar{\alpha}) \approx 0.2573$ when $d = 8$. This local limit theorem (in the range of moderate deviations) also implies the following convergence rate

$$\sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} < x\right) - \Phi(x)\right| = \begin{cases} O(n^{-1/2}), & \text{if } 1 \leq d \leq 7;\\ O(n^{-3(3/2-\sqrt{2})}), & \text{if } d = 8, \end{cases} \tag{62}$$

where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-t^2/2}\,\mathrm{d}t$.

Moment generating function of X_n normalized by that of a normal distribution with the same mean and variance. Let $\Pi_n(y) := \mathbb{E}(e^{X_n y})$ and $\phi_n(y) := e^{-\mathbb{E}(X_n)y-\mathbb{V}(X_n)y^2/2}\Pi_n(y)$. From the recurrence (5), we have

$$\phi_n(y) = \sum_{j_1+\cdots+j_{2d}=n-1} \pi_{n,\mathbf{j}}\phi_{j_1}(y)\cdots\phi_{j_{2d}}(y)e^{\Delta_{n,\mathbf{j}}y+\nabla_{n,\mathbf{j}}y^2} \qquad (n \geq 1),$$

with $\phi_0(y) = 1$, where

$$\Delta_{n,\mathbf{j}} = \delta_{n,1} + \mathbb{E}(X_{j_1}) + \cdots + \mathbb{E}(X_{j_{2d}}) - \mathbb{E}(X_n),$$
$$\nabla_{n,\mathbf{j}} = \frac{1}{2}\left(\mathbb{V}(X_{j_1}) + \cdots + \mathbb{V}(X_{j_{2d}}) - \mathbb{V}(X_n)\right).$$

Note that $\phi_n(y)$ is in general not a moment generating function.

Recurrences. Define $\phi_{n,k} := \phi_n^{(k)}(0)$. Then by the recurrence of $\phi_n(y)$, we have

$$\phi_{n,k} = \psi_{n,k} + 2^d \sum_{0\leq j<n} \pi_{n,j}\phi_{j,k} \qquad (n \geq 1),$$

where $\phi_{0,k} = 0$ and

$$\psi_{n,k} = \sum_{\substack{i_0+i_1+\cdots+i_{2d}+2i_{2d+1}=k \\ 0\leq i_1,\ldots,i_{2d}<k}} \frac{k!}{i_0!\cdots i_{2^d}!i_{2^d+1}!} \sum_{j_1+\cdots+j_{2^d}=n-1} \pi_{n,\mathbf{j}}\phi_{j_1,i_1}\cdots\phi_{j_{2d},i_{2d}}\Delta_{n,\mathbf{j}}^{i_0}\nabla_{n,\mathbf{j}}^{i_{2^d+1}}.$$

A uniform upper bound for $\phi_{n,k}$. Recall that $\bar{\alpha} = 1/3$ when $1 \leq d \leq 7$, and $\bar{\alpha} = \sqrt{2}-1$ when $d = 8$. We will prove, by an inductive argument, that

$$|\phi_{n,k}| \leq k!A^k n^{k\bar{\alpha}} \qquad (k, n \geq 0), \tag{63}$$

where A is a suitable constant that will be specified later. Note that (63) holds for $k = 0, 1, 2$.

An upper bound for $\Delta_{n,\mathbf{j}}$. By the estimate (49), we have

$$\Delta_{n,\mathbf{j}} = O\left(n^{\alpha}\right) = \begin{cases} O\left(n^{1/3-\varepsilon}\right), & \text{if } 1 \leq d \leq 7; \\ O\left(n^{\sqrt{2}-1}\right), & \text{if } d = 8, \end{cases} \tag{64}$$

uniformly for all tuples $(j_1, \ldots, j_{2^d})$.

An upper bound for $\nabla_{n,\mathbf{j}}$. We need to refine the asymptotic estimate (4). Since the variance satisfies the recurrence

$$\mathbb{V}(X_n) = \sum_{j_1+\cdots+j_{2d}=n-1} \pi_{n,\mathbf{j}}\Delta_{n,\mathbf{j}}^2 + 2^d \sum_{0\leq j<n} \pi_{n,j}\mathbb{V}(X_j),$$

and the first sum on the right-hand side is bounded above by

$$\sum_{j_1+\cdots+j_{2d}=n-1} \pi_{n,\mathbf{j}}\Delta_{n,\mathbf{j}}^2 = \begin{cases} O\left(n^{2/3-2\varepsilon}\right), & \text{if } 1 \leq d \leq 7; \\ O\left(n^{2\sqrt{2}-2}\right), & \text{if } d = 8, \end{cases}$$

we obtain, by applying Corollary (1),

$$\mathbb{V}(X_n) = \sigma_d^2 n + \begin{cases} O\left(n^{2/3-2\varepsilon}\right), & \text{if } 1 \le d \le 7; \\ O\left(n^{2(\sqrt{2}-1)}\right), & \text{if } d = 8. \end{cases}$$

This implies that

$$\nabla_{n,\mathbf{j}} = \begin{cases} O\left(n^{2/3-2\varepsilon}\right), & \text{if } 1 \le d \le 7; \\ O\left(n^{2(\sqrt{2}-1)}\right), & \text{if } d = 8. \end{cases} \tag{65}$$

An estimate for $\phi_{n,3}$. From (64) and (65), it follows that

$$\psi_{n,3} = \begin{cases} O\left(n^{1-\varepsilon}\right), & \text{if } 1 \le d \le 7; \\ O\left(n^{3(\sqrt{2}-1)}\right), & \text{if } d = 8. \end{cases}$$

Thus (63) holds for $k = 3$ by applying (12) when $1 \le d \le 7$ and (15) when $d = 8$.

Induction. For higher values of k, we use the estimates (by (64) and (65))

$$|\Delta_{n,\mathbf{j}}| \le K_5 n^{\bar{\alpha}}, \quad |\nabla_{n,\mathbf{j}}| \le K_6 n^{2\bar{\alpha}}, \tag{66}$$

uniformly for all tuples $(j_1, \dots, j_{2^d})$.

Assume that (63) holds $\phi_{n,i}$ for $i < k$. Then by (66) and induction

$$\begin{aligned} |\psi_{n,k}| &\le k! n^{k\bar{\alpha}} \sum_{\substack{i_0+\cdots+i_{2^d}+2i_{2^d+1}=k \\ 0 \le i_1,\dots,i_{2^d}<k}} A^{i_1+\cdots+i_{2^d}} \frac{K_5^{i_0} K_6^{i_{2^d+1}}}{i_0! i_{2^d+1}!} \sum_{j_1+\cdots+j_{2^d}=n-1} \pi_{n,\mathbf{j}} \left(\frac{j_1}{n}\right)^{i_1\bar{\alpha}} \cdots \left(\frac{j_{2^d}}{n}\right)^{i_{2^d}\bar{\alpha}} \\ &\le k! n^{k\bar{\alpha}} e^{K_5+K_6} \sum_{0 \le \ell \le k} A^\ell S(\ell), \end{aligned} \tag{67}$$

where

$$S(\ell) := \sum_{i_1+\cdots+i_{2^d}=\ell} \sum_{j_1+\cdots+j_{2^d}=n-1} \pi_{n,\mathbf{j}} \left(\frac{j_1}{n}\right)^{i_1\bar{\alpha}} \cdots \left(\frac{j_{2^d}}{n}\right)^{i_{2^d}\bar{\alpha}}.$$

An estimate for $S(\ell)$. We now show that $S(\ell) \to 0$ as $\ell \to \infty$.

Lemma 5. *For $\ell \ge 0$*

$$S(\ell) \le c(\ell\bar{\alpha}+1)^{-d} \qquad (d \ge 1), \tag{68}$$

where $c > 0$ is independent of ℓ and n.

Proof. First, by the strong law of large numbers

$$\begin{aligned} S(\ell) &\le c \int_{[0,1]^d} \sum_{i_1+\cdots+i_{2^d}=\ell} \prod_{1 \le h \le 2^d} q_h(\mathbf{x})^{i_h\bar{\alpha}} \mathrm{d}\mathbf{x} \\ &= c 2^d [z^\ell] \int_{[0,1/2]^d} \prod_{1 \le h \le 2^d} \frac{1}{1 - q_h(\mathbf{x})^{\bar{\alpha}} z} \mathrm{d}\mathbf{x}. \end{aligned}$$

Observe that the smallest term among the $q_h(\mathbf{x})$'s is $q_{2^d}(\mathbf{x}) = (1-x_1)\cdots(1-x_d)$ when $\mathbf{x} \in [0,1/2]^d$. Thus the dominant term for large ℓ comes from $q_{2^d}(\mathbf{x})$, and it follows that

$$\begin{aligned}
[z^\ell]\int_{[0,1/2]^d}\prod_{1\le h\le 2^d}\frac{1}{1-q_h(\mathbf{x})^{\bar\alpha}z}\,\mathrm{d}\mathbf{x} &\sim \int_{[0,1/2]^d} q_{2^d}(\mathbf{x})^{\ell\bar\alpha}\prod_{1\le h<2^d}\frac{1}{1-q_h(\mathbf{x})/q_{2^d}(\mathbf{x})}\,\mathrm{d}\mathbf{x}\\
&\sim \int_{[0,1/2]^d}(1-x_1)^{\ell\bar\alpha}\cdots(1-x_1)^{\ell\bar\alpha}\,\mathrm{d}\mathbf{x}\\
&\sim (\ell\bar\alpha+1)^{-d}.
\end{aligned}$$

This proves (68). ∎

Proof of (63). Substituting the estimate (68) into (67), we obtain

$$|\psi_{n,k}| \le \frac{c}{(k\bar\alpha+1)^d}k!A^k n^{k\bar\alpha}.$$

Then, by the asymptotic transfer (15),

$$|\phi_{n,k}| \le \frac{c'}{(k\bar\alpha+1)^d}k!A^k n^{k\bar\alpha},$$

where c' is independent of n and k. Thus $c'/(k\bar\alpha+1)^d < 1$ for large enough k, say $k \ge k_0$. Hence, (63) follows by suitably tuning A for $k \le k_0$; see [1] for similar details.

An estimate for the characteristic function for small y. Denote by $\varphi_n(y) = \Pi_n(iy/\sqrt{\mathbb{V}(X_n)})$. Then, by (63) and the Taylor series expansion,

$$\left|\varphi_n(y) - e^{-y^2/2}\right| = O\left(|y|^3 n^{-3(1/2-\bar\alpha)}e^{-y^2/2}\right) \tag{69}$$

for $|y| \le \varepsilon_0 n^{1/2-\bar\alpha}$, where $\varepsilon_0 > 0$ is sufficiently small.

A uniform estimate for $\Pi_n(iy)$ for $|y| \le \varepsilon$. From (69), we deduce that

$$|\Pi_n(iy)| \le e^{-\varepsilon_1(n+1)y^2} \qquad (n \ge 3), \tag{70}$$

for $|y| \le \varepsilon_0 n^{-\bar\alpha}$, where ε_1 is a suitably chosen small constant.

We now prove that the estimate (70) indeed holds for $|y| \le \varepsilon_2$, $\varepsilon_2 > 0$ being a small constant. To that purpose, choose n_0 large enough and set $\varepsilon_2 := \varepsilon_0 n_0^{-\bar\alpha}$. Then, (70) holds for $3 \le n \le n_0$ and $|y| \le \varepsilon_2$. For $n > n_0$, by (5) and induction,

$$\begin{aligned}
|\Pi_n(iy)| &\le \sum_{j_1+\cdots+j_{2^d}=n-1}\pi_{n,\mathbf{j}}|\Pi_{j_1}(iy)|\cdots|\Pi_{j_{2^d}}(iy)|\\
&\le e^{-\varepsilon_1(n+1)y^2-\varepsilon_1(2^d-2)y^2}\\
&\le e^{-\varepsilon_1(n+1)y^2}.
\end{aligned}$$

This concludes the induction proof.

Reformulating the estimate (70) yields the following global estimate for $\varphi_n(y)$

$$|\varphi_n(y)| = O\left(e^{-\varepsilon n y^2}\right) \qquad (n \ge 3), \tag{71}$$

uniformly for $|y| \le \varepsilon_2 n^{1/2}$.

Berry-Esseen bounds and local limit theorems. The convergence rates (62) now follows by (69), (71) and the Berry-Esseen smoothing inequality

$$\sup_x \left| \mathbb{P}\left(\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} < x \right) - \Phi(x) \right| = O\left(R_n^{-1} + \int_{R_n}^{R_n} \left| \frac{\varphi_n(y) - e^{-y^2/2}}{y} \right| \mathrm{d}t \right),$$

where $R_n := \varepsilon n^{3(1/2-\bar{\alpha})}$; see [42].

For local limit theorems, we first observe that the span of X_n is 1 by induction, so that (70) can be extended to $|y| \leq \pi$ (again by induction). Then Theorem 4 follows by applying the Fourier inversion formula

$$\mathbb{P}(X_n = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iky} \Pi_n(iy)\,\mathrm{d}y,$$

where $k = \left\lfloor \mathbb{E}(X_n) + x\sqrt{\mathbb{V}(X_n)} \right\rfloor$; see Figure 3.

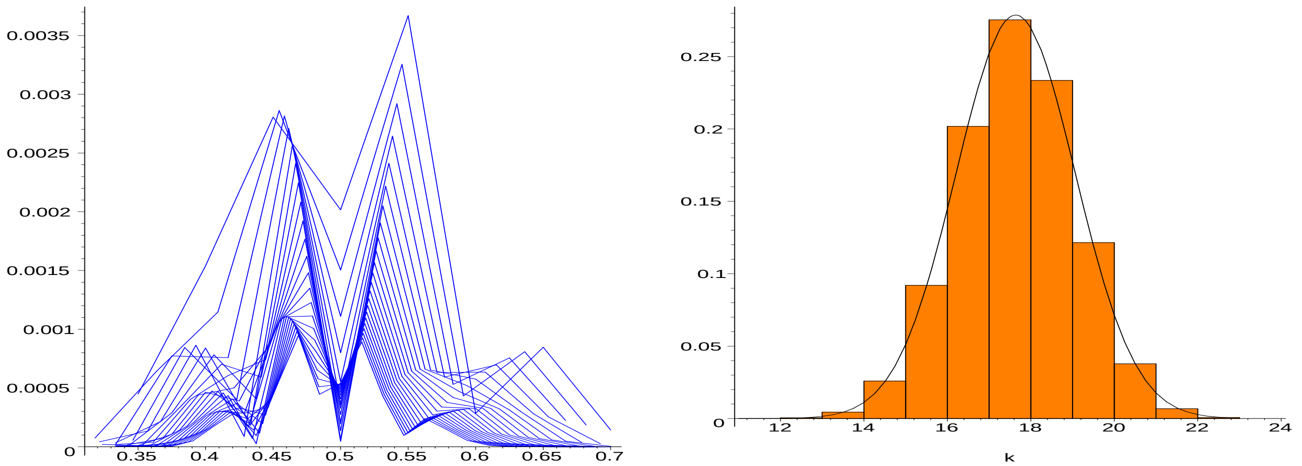

Figure 3: *Left: A Sedgewick plot of the absolute difference between* $\mathbb{P}(X_n = k)$ *and* $e^{-(k-\mathbb{E}(X_n))^2/(2\mathbb{V}(X_n))}/\sqrt{2\pi\mathbb{V}(X_n)}$ *for* $n = 20, 22, \dots, 64$ *and* $\lfloor 0.35n \rfloor \leq k \leq \lfloor 0.7n \rfloor$ *(normalized in the unit interval) when* $d = 2$. *Right: the histogram of* $\mathbb{P}(X_n = k)$ *for* $d = 3, n = 30$ *and* $k = 12, \dots, 23$, *together with the corresponding normal curve (having the same mean and variance).*

Extensions to general cost measures. The same method of proof applies to other cost measures in random quadtrees. In particular, Assume that T_n in (57) is deterministic and satisfies $T_n = O(n^\rho)$, where $\rho < 1/2$. If $1 \leq d \leq 7$, then we have the following Berry-Esseen bounds for Y_n.

$$\sup_x \left| \mathbb{P}\left(\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} < x \right) - \Phi(x) \right| = \begin{cases} O(n^{-1/2}), & \text{if } \rho < 1/3; \\ O(n^{-1/2}\log n), & \text{if } \rho = 1/3; \\ O(n^{-3(1/2-\rho)}), & \text{if } 1/3 < \rho < 1/2. \end{cases}$$

When $d = 8$, then

$$\sup_x \left| \mathbb{P}\left(\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} < x \right) - \Phi(x) \right|$$
$$= \begin{cases} O(n^{-3(3/2-\sqrt{2})}), & \text{if } \rho < \sqrt{2}-1; \\ O(n^{-3(3/2-\sqrt{2})}(\log n)^3), & \text{if } \rho = \sqrt{2}-1; \\ O(n^{-3(1/2-\rho)}), & \text{if } \sqrt{2}-1 < \rho < 1/2. \end{cases}$$

The corresponding local limit theorems can be derived when Y_n assumes only integer values.

5 Random d-dimensional grid-trees

We consider briefly the phase changes in random grid-trees in this section, the required asymptotic transfers being also given.

Grid trees. Devroye [12] extended the d-dimensional point quadtrees and m-ary search trees as follows. Instead of choosing the first point as the root, one chooses, say the first $m-1$ points ($m \geq 2$) and places them at the root. These $m-1$ points then split the space into m^d smaller regions (called grids) when no pair of points is collinear. Each node in the corresponding grid-tree has at most m^d subtrees. When $m = 2$, grid-trees are quadtrees; when $d = 1$, grid-trees reduce to the usual m-ary search trees; see [37].

Random grid-trees. Fix $m \geq 2$ and $d \geq 1$ throughout this section. Assume that the input is a sequence of n random points uniformly and independently chosen from $[0,1]^d$. Construct the grid-tree from this sequence. The resulting tree is called a *random grid-tree*.

Phase changes of the number of leaves. For simplicity of presentation, we consider the number of leaves in random grid-trees, denoted by X_n.

m	2	3	4	$5,\dots,8$	$9,\dots,\underline{26}$
d	$1,\dots,\underline{8}$	$1,\dots,4$	$1,\dots,3$	$1,2$	1

Table 4: The set $\mathcal{S}$ of all pairs of (m,d) for which X_n is asymptotically normally distributed. The two boundary cases $(2,26)$ (m-ary search trees) and $(1,8)$ (quadtrees) are both underlined.

Theorem 5. *If $(m,d) \in \mathcal{S}$, where $\mathcal{S}$ is given in Table 4, then*

$$\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \xrightarrow{\mathscr{M}} N(0,1);$$

if $m \geq 2, d \geq 1$ and $(m,d) \notin \mathcal{S}$, then the sequence of random variables $(X_n - \mathbb{E}(X_n))/\sqrt{\mathbb{V}(X_n)}$ does not converge to a fixed limit law.

More refined results (and more phase changes) can be derived as in the case of quadtrees.

Recurrence of X_n. The recurrence of X_n now has the form

$$X_n \stackrel{\mathscr{D}}{=} \sum_{1\le j\le m^d} X_{J_j}^{(j)} + \delta_{n,1}, \qquad (n\ge 1),$$

with $X_0 = 0$, where $X_n, X_n^{(1)}, \dots, X_n^{(m^d)}, (J_1,\dots,J_{m^d})$ are independent and $X_n \stackrel{\mathscr{D}}{=} X_n^{(j)}, 1\le j\le m^d$. Moreover, the splitting probabilities can be expressed as

$$\begin{aligned}\pi_{n,\mathbf{j}} &= \mathbb{P}\left(J_1 = j_1, \dots, J_{m^d} = j_{m^d}\right)\\ &= \binom{n-m+1}{j_1,\dots,j_{m^d}} \int_{([0,1]^d)^{m-1}} \prod_{\substack{1\le h\le m^d\\ h-1=(b_1,\dots,b_d)_m}} q_h(\mathbf{x}_1,\dots,\mathbf{x}_{m-1})^{j_h} \mathrm{d}\mathbf{x}_1\dots \mathrm{d}\mathbf{x}_{m-1}\end{aligned}$$

for all $j_1+\dots+j_{m^d} = n-m+1$, where

$$q_h(\mathbf{x}_1,\dots,\mathbf{x}_{m-1}) = \prod_{1\le i\le d}\sum_{0\le \ell<m} 1_{\{\ell\}}(b_i)\left(x_{(\ell+1)}^{(i)} - x_{(\ell)}^{(i)}\right),$$

with $x_{(\ell)}$ denoting the ℓ-th order statistic of $x_1,\dots,x_{m-1}$ ($x_{(0)} := 0, x_m := 1$).

Recurrence of moments. All moments satisfy recurrences of the form

$$A_n = B_n + m^d \sum_{0\le j\le n-m+1} \pi_{n,j} A_j, \qquad (n\ge m-1), \tag{72}$$

where $\pi_{n,j}$ denotes the probability that a specified subtree (say the first) of the root has j nodes.

We now show that $\pi_{n,j}$ can be expressed in the form

$$\pi_{n,j} = \sum_{j\le j_1\le\dots\le j_{d-1}\le n-m+1} \frac{\binom{n-1-j_{d-1}}{m-2}}{\binom{n}{m-1}} \prod_{1\le i<d} \frac{\binom{j_i - j_{i-1}+m-2}{m-2}}{\binom{j_i+m-1}{m-1}}. \tag{73}$$

To that purpose, we first split $\pi_{n,j}$ as follows.

$$\pi_{n,j} = \sum_{j\le i_1\le i_2\le\dots\le i_{d-1}\le n-m+1} \varpi_{j;i_1,\dots,i_{d-1}},$$

where $\varpi_{j;i_1,\dots,i_{d-1}}$ denotes the probability that the n random points are distributed in the d-dimensional unit cube in the following way: the first $m-1$ points, denoted by $\mathbf{x}_1,\dots,\mathbf{x}_{m-1}$, split $[0,1]^d$ into m^d grids and the remaining points are placed in these grids such that grids of the form

$$\left[0, x_{(1)}^{(1)}\right]\times\dots\times\left[0, x_{(1)}^{(i)}\right]\times\left[x_{(1)}^{(i+1)}, 1\right] \qquad (i = 0,\cdots,d),$$

contain $n-m-i_{d-1}+1, i_{d-1}-i_{d-2},\dots,i_1-j, j$ random points, respectively.

By definition, we have

$$\begin{aligned}\frac{\varpi_{j;i_1,\dots,i_{d-1}}}{\binom{n-m+1}{i_0,i_1-i_0,\dots,i_d-i_{d-1}}} &= \int_{([0,1]^d)^{m-1}} \prod_{1\le i\le d}\left(x_{(1)}^{(i)}\right)^{i_{d-i}}\left(1-x_{(1)}^{(i)}\right)^{i_{d-i+1}-i_{d-i}} \mathrm{d}\mathbf{x}_1\dots\mathrm{d}\mathbf{x}_{m-1}\\ &= \prod_{1\le r\le d}\int_{[0,1]^{m-1}} \left(x_{(1)}^{(r)}\right)^{i_{d-r}}\left(1-x_{(1)}^{(r)}\right)^{i_{d-r+1}-i_{d-r}} \mathrm{d}x_1^{(r)}\dots\mathrm{d}x_{m-1}^{(r)},\end{aligned} \tag{74}$$

where $i_0 := j$ and $i_d := n - m + 1$. It remains to evaluate integrals of the form

$$\int_{[0,1]^{m-1}} x_{(1)}^{\rho} \left(1 - x_{(1)}\right)^{\tau} \mathrm{d}x_1 \ldots \mathrm{d}x_{m-1},$$

where $\rho, \tau \geq 0$. By dividing the domain of integration into $(m-1)!$ sets of the form $\{(x_1, \ldots, x_{m-1}) | x_{\sigma(1)} < \cdots < x_{\sigma(m-1)}\}$, where σ runs through all permutations of $m-1$ elements

$$\begin{aligned}\int_{[0,1]^{m-1}} x_{(1)}^{\rho} \left(1 - x_{(1)}\right)^{\tau} \mathrm{d}x_1 \ldots \mathrm{d}x_{m-1} &= (m-1)! \int_{0 \leq x_1 \leq \cdots \leq x_{m-1} \leq 1} x_1^{\rho} (1 - x_1)^{\tau} \mathrm{d}x_1 \ldots \mathrm{d}x_{m-1} \\ &= (m-1) \int_0^1 x_1^{\rho} (1 - x_1)^{\beta+m-2} \mathrm{d}x_1 \\ &= (m-1) \frac{\Gamma(\rho+1)\Gamma(\tau+m-1)}{\Gamma(\rho+\tau+m)},\end{aligned}$$

by symmetry. Substituting this expression into (74) gives the desired result (73). ∎

The DE. Let $A(z) = \sum_{n\geq 0} A_n z^n$, $B(z) = \sum_{n\geq 1} B_n z^n$, and $f = A - B$. Then the recurrence (72) translates into the DE

$$(1-z)^{m-1}\mathbb{D}^{m-1} \left(z^{m-1}(1-z)^{m-1}\mathbb{D}^{m-1}\right)^{d-1} f(z) = m!^d A(z),$$

or, in terms of the ϑ-operator,

$$\vartheta^{\overline{m-1}} \left(z^{m-1}\vartheta^{\overline{m-1}}\right)^{d-1} f(z) = m!^d A(z), \tag{75}$$

where $\vartheta^{\overline{m-1}} = \vartheta(\vartheta+1)\cdots(\vartheta+m-2)$.

The normal form. We then rewrite the DE in the form

$$P_0(\vartheta) f(z) = \sum_{1\leq j\leq (m-1)(d-1)} (1-z)^j P_j(\vartheta) f(z) + m!^d B(z),$$

where the P_j's are polynomials of degree dm. In particular,

$$P_0(\vartheta) = (\vartheta^{\overline{m-1}})^d - m!^d = \prod_{1\leq j\leq d} \left(\vartheta^{\overline{m-1}} - m! e^{2j\pi i/d}\right).$$

The unique case when the above DE reduces to a pure Cauchy-Euler type is $d = 1$. Also the "linearization" achieved by the Euler transform does not seem to work directly for $m \geq 3$. This says that it is not obvious how to derive an explicit expression such as (38) when $m \geq 3$.

Zeros of $P_0(x)$. Our method of proof for deriving the asymptotic transfers is mostly operational and requires only limited properties of the zeros of the indicial polynomial $P_0(x)$. The proofs of the following properties are straightforward and thus omitted.

- The zero with the largest real part is $x = 2$. All other zeros have real parts strictly less than 2.
- All zeros of $P_0(x)$ are simple (we need only this property for $x = 2$ and the second largest zeros in real part).

Other properties similar to those for the case $d = 1$ (m-ary search trees) can be derived as in [37, Ch. 3].

Asymptotic transfers. We state the main asymptotic transfers needed for proving Theorem 5.

Let $H_m := \sum_{1\le j\le m} 1/j$ denotes the harmonic numbers. Define

$$K_B := \frac{1}{d(H_m-1)} \sum_{k\ge 0} V_k B^*(k+2), \tag{76}$$

when the series converges, where V_k is defined recursively by $V_k = 0$ when $k < 0$, $V_0 = 1$, and

$$V_k = \sum_{1\le \ell\le (m-1)(d-1)} \frac{P_\ell(k+2)}{P_0(k+2)} V_{k-\ell} \qquad (k \ge 1),$$

and $B^*(s) := \int_0^1 B(x)(1-x)^{s-1}\,\mathrm{d}x$ when the integral converges.

Theorem 6. *Let A_n be defined by the recurrence (72) with A_0 and $\{B_n\}_{n\ge 1}$ given. Then*

(i) *(Small toll functions)*

$$A_n \sim K_B n \quad \textit{iff} \quad B_n = o(n) \;\; \textit{and} \;\; \left|\sum_n B_n n^{-2}\right| < \infty,$$

where the constant K_B is given in (76);

(ii) *(Linear toll functions) Assume that $B_n = cn + u_n$, where $c \in \mathbb{C}$ and u_n is a sequence of complex numbers. Then*

$$A_n \sim \frac{c}{d(H_m-1)} n \log n + K_1 n \quad \textit{iff} \quad u_n = o(n) \; \textit{and} \left|\sum_n u_n n^{-2}\right| < \infty.$$

Here $K_1 := cK_2 + K_u$ with K_u defined by replacing the sequence B_n by u_n in (76) and K_2 given explicitly by

$$K_2 := \frac{1}{d(H_m-1)} \left(\sum_{k\ge 1} \frac{V_k}{k(k+1)} + \gamma - 2 - \frac{d}{2}(H_m-1) + \frac{H_m^{(2)}-1}{2(H_m-1)} \right),$$

where $H_m^{(2)} := \sum_{1\le j\le m} 1/j^2$.

(iii) *(Large toll functions) Assume that $\Re(\upsilon) > 1$ and $c \in \mathbb{C}$. Then*

$$B_n \sim cn^{\upsilon} \quad \textit{iff} \quad A_n \sim \frac{c((\upsilon+1)^{\overline{m-1}})^d}{((\upsilon+1)^{\overline{m-1}})^d - m!^d} n^{\upsilon}.$$

In particular, if $d = 1$, then $V_k = \delta_{k,0}$ and

$$K_B = \frac{B^*(2)}{H_m-1} = \frac{1}{H_m-1} \sum_{k\ge 0} \frac{B_k}{(k+1)(k+2)};$$

see [3].

Growth order of V_k for grid-trees. The sequence V_k satisfies the DE

$$\begin{aligned}&\left((\mathbb{D}_z z+m-2)\cdots(\mathbb{D}_z z+1)\mathbb{D}_z z(1-z)^{m-1}\right)^{d-1}\\ &\qquad\times(\mathbb{D}_z z+m-2)\cdots(\mathbb{D}_z z+1)\mathbb{D}_z\left(z^2V(z)\right)-m!^d zV(z)=0,\end{aligned}$$

implying that the solution of the form $V(z)=(1-z)^{-s}\phi(1-z)$ has the indicial equation

$$s^d(s+1)^d\cdots(s+m-2)^d=0.$$

Thus we deduce that

$$V_k=O\left(k^{-1}(\log k)^c\right),$$

for some $c\geq d-2$. This implies that the series in (76) is convergent for both cases of small and linear toll functions.

Refining the asymptotic transfer for small toll functions. To derive the second-order term for $\mathbb{E}(X_n)$ and $\mathbb{V}(X_n)$, we also need the following types of transfer.

Let $\alpha+1$ denote the real part of the second largest zeros of $P_0(x)$ (all zeros arranged in decreasing order according to their real parts), and $\beta>0$ denote the absolute value of the imaginary part of either zero.

Proposition 2. *Assume that A_n satisfies (72).*

(i) *If $B_n\sim cn^{\upsilon}$, where $c\in\mathbb{C}$ and $\alpha<\Re(\upsilon)<1$, then*

$$A_n=K_Bn+\frac{c((\upsilon+1)^{\overline{m-1}})^d}{((\upsilon+1)^{\overline{m-1}})^d-m!^d}n^{\upsilon}+o(n^{\upsilon}+n^{\varepsilon}),$$

where K_B is defined in (76).

(ii) *If $B_n=o(n^{\alpha})$, then*

$$A_n=K_Bn+K(\lambda_1)n^{\alpha+i\beta}+K(\lambda_2)n^{\alpha-i\beta}+o(n^{\alpha}+n^{\varepsilon}),$$

where the $K(\lambda_j)$'s are constants whose expressions are similarly defined as in (48). If the B_k's are all real, then $K(\lambda_1)=\overline{K(\lambda_2)}$.

These types of transfer and the inductive arguments used for quadtrees can be applied to prove local limit theorems for X_n with optimal convergence rates. Limit theorems for many other shape parameters can also be derived. We mention only the application to total path length.

Total path length. Neininger and Rüschendorf [40] derived a general limit law for the total path length in random split trees of Devroye (see [12]), which cover in particular grid-trees. Their result is based on the assumption that the expected total path length satisfies asymptotically $cn\log n+c'n$. Our asymptotic transfer for linear toll functions shows that this is the case for grid-trees. This proves the limit law for the total path length in random grid-trees. Note that the limit law can also be derived directly by method of moments and our asymptotic transfer for large toll functions.

6 Conclusions

We extended in this paper the asymptotic theory for Cauchy-Euler DEs developed in [7] to essentially DEs with polynomial coefficients (often referred to as *holonomic DEs*) and $z = 0$ not an irregular singularity. Not only the results are very general, but also the method of proof requires almost no knowledge on DEs. Indeed, since all our manipulations are based on linear operators, only properties of the first-order DEs are used, which can be further avoided by completely operating on recurrences of quicksort type (see [30]). The main feature of such an approach is that all differential operators are regarded as coefficient-transformers, so that no analytic properties are needed for the functions involved.

We applied the general asymptotic transfers developed in this paper to clarify the phase changes of limit laws in quadtrees and more general grid-trees. Further applications to distributional properties of profiles of random search trees will be given elsewhere.

For more methodological interest, we conclude this paper by mentioning an alternative approach to proving general asymptotic transfers for A_n (under suitable growth information on B_n) based solely on the theory of differential equations. Such an approach was inspired by the series of papers by Flajolet and his coauthors (see [17, 20, 22, 26]). We start from the method of Frobenius and seeks solutions of the form $(1-z)^{-\lambda_k}\phi(1-z)$ for the homogeneous DE $(\vartheta(z\vartheta)^{d-1} - 2^d)f(z) = 0$, where $\phi(z)$ is analytic at $z = 0$. A detailed information on the zeros of $P_0(x)$ is needed; in particular, we can show that when d is a multiple of 6 there are two pairs of non-real zeros differing by integers (in that case, logarithmic terms need to be introduced). Then we use the method of variation of parameters (see [32]) for the non-homogeneous DE; a long and laborious calculation of the Wronskians then leads to the form

$$\begin{aligned} f(z) = &\sum_{0\le j<d} \xi_j(z)(1-z)^{-\lambda_j} \\ &+ 2^d \sum_{0\le j<d} \eta_j(z)(1-z)^{-\lambda_j} \int_0^z (1-t)^{\lambda_j-1} B(t) \sum_{0\le r\le\kappa_d} \zeta_{j,r}(t) \left(\log\frac{z}{t}\right)^r \mathrm{d}t, \end{aligned} \tag{77}$$

where $\kappa_d \le (d-1)^2$ and $\xi_j, \eta_j, \zeta_{j,r}$ are functions analytic in the unit circle satisfying $\sum_n |[z^n]\chi(z)| < \infty$, where $\chi \in \{\xi_j, \eta_j, \zeta_{j,r}\}$. Similar expressions can be derived for $\sum_{1\le j<d}(1-z)^j P_j(\vartheta) f$. Then the sufficiency proofs of the transfers (12), (13), (15) are reduced to deriving asymptotic transfers for integrals of the form

$$\xi(z)(1-z)^{-\upsilon} \int_0^z (1-t)^{\upsilon-1} B(t)\eta(t) \left(\log\frac{z}{t}\right)^r \mathrm{d}t.$$

Such a general approach, although quickly gives the general form of the solution, does not seem easily amended for getting expressions for the leading constants (similar to most asymptotic problems on DEs and linear differential systems); also for more general DEs such as (75), the precise characterizetion of the zero locations (of their differences) requires more delicate analysis.

Acknowledgements

We thank Philippe Flajolet for showing us the phase change phenomena in random fragmentation processes and suggesting the current study. Part of the work of the third author was carried out while he was visiting Institut für Stochastik und Mathematische Informatik, J. W. Goethe-Universität (Frankfurt); he thanks the Institute for its hospitality and support.

References

[1] Z.-D. Bai, H.-K. Hwang and T.-H. Tsai (2003). Berry-Esseen bounds for the number of maxima in planar regions. *Electronic Journal of Probability* **8** Article No. 9, 26 pp.

[2] B. Chauvin and N. Pouyanne (2004). m-ary search trees when $m \geq 27$: a strong asymptotics for the space requirements. *Random Structures and Algorithms* **24** 133–154.

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

[4] H.-H. Chern and H.-K. Hwang (2003). Partial match queries in random quadtrees. *SIAM Journal on Computing* **32** 904–915.

[5] H.-H. Chern and H.-K. Hwang (2003). Partial match queries in random k-d trees. Manuscript submitted for publication.

[6] H.-H. Chern and H.-K. Hwang (2004). Limit distribution of the number of consecutive records. *Random Structures and Algorithms*, accepted for publication.

[7] 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.

[8] D. S. Dean and S. N. Majumdar (2002). Phase transition in a random fragmentation problem with applications to computer science. *Journal of Physics. A. Mathematical and General* **35** L501–L507.

[9] M. de Berg, M. van Kreveld, M. Overmars, O. Schwarzkopf (2000). *Computational Geometry: Algorithms and Applications*. Second revised edition, Springer-Verlag.

[10] L. Devroye (1987). Branching processes in the analysis of the heights of trees. *Acta Informatica* **24** 277–298.

[11] L. Devroye (1991). Limit laws for local counters in random binary search trees. *Random Structures and Algorithms* **2** 303–315.

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

[13] L. Devroye and L. Laforest (1990). An analysis of random d-dimensional quad trees. *SIAM Journal on Computing* **19** 821–832.

[14] A. Erdélyi, W. Magnus, F. Oberhettinger and F. Tricomi (1953). *Higher Transcendental Functions*. Volume I. Robert E. Krieger Publishing Co., Fla.

[15] J. A. Fill and N. Kapur (2004). The space requirement of m-ary search trees: distributional asymptotics for $m \geq 27$. In *Proceedings of the 7th Iranian Statistical Conference*.

[16] R. A. Finkel and J. L. Bentley (1974). Quadtrees: A data structure for retrieval on composite keys. *Acta Informatica* **4** 1–9.

[17] P. Flajolet, G. Gonnet, C. Puech and J. M. Robson (1993). Analytic variations on quadtrees. *Algorithmica* **10**, 473–500.

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

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

[20] P. Flajolet and T. Lafforgue (1994). Search costs in quadtrees and singularity perturbation asymptotics. *Discrete and Computational Geometry* **12** 151–175.

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

[22] P. Flajolet and P. Puech (1986). Partial match retrieval of multi-dimensional data. *Journal of the ACM* **32** 371–407.

[23] P. Flajolet and R. Sedgewick (1995). Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theoretical Computer Science* **144**, 101–124.

[24] P. Flajolet and R. Sedgewick (1995). *Analytic Combinatorics*. Book in preparation; current version available via the link algo.inria.fr/flajolet/Publications/AnaCombi1to9.pdf.

[25] D. Foata, G.-N. Han and B. Lass (2001). Les nombres hyperharmoniques et la fratrie du collectionneur de vignettes. *Seminaire Lotharingien de Combinatoire* **47** Article B47a.

[26] M. Hoshi and P. Flajolet (1992). Page usage in a quadtree index. *BIT* **32** 384–402.

[27] H.-K. Hwang (1998). On convergence rates in the central limit theorems for combinatorial structures. *European Journal of Combinatorics,* **19** 329–343.

[28] H.-K. Hwang (2003). Second phase changes in random m-ary search trees and generalized quicksort: convergence rates. *Annals of Probability* **31** 609–629.

[29] H.-K. Hwang (2004). Phase changes in random recursive structures and algorithms (a brief survey). In *Proceedings of the Workshop on Probability with Applications to Finance and Insurance*. Edited by T. L. Lai, H. Yang and S. P. Yung, World Scientific, pp. 82–97.

[30] H.-K. Hwang and R. Neininger (2002). Phase change of limit laws in the quicksort recurrence under varying toll functions. *SIAM Journal on Computing* **31** 1687–1722.

[31] H.-K. Hwang and T.-H. Tsai (2003). An asymptotic theory for recurrence relations based on minimization and maximization. *Theoretical Computer Science* **290** 1475–1501.

[32] E. L. Ince (1926). *Ordinary Differential Equations,* Dover, New York.

[33] S. Janson (2004). Functional limit theorems for multitype branching processes and generalized Polya urns. *Stochastic Processes and their Applications* **110** 177–245.

[34] G. Labelle and L. Laforest (1995). Combinatorial variations on multidimensional quadtrees. *Journal of Combinatorial Theory, Series A* **69** 1–16.

[35] G. Labelle and L. Laforest (1995). Sur la distribution de l'arité de la racine d'une arborescence hyperquaternaire à d dimensions. *Discrete Mathematics* **139** 287–302.

[36] G. Labelle and L. Laforest (1996). Étude de constantes universelles pour les arborescences hyper-quaternaires de recherche. *Discrete Mathematics* **153** 199–211.

[37] H. M. Mahmoud (1992). *Evolution of Random Search Trees*. John Wiley & Sons, New York.

[38] C. Martínez, A. Panholzer and H. Prodinger (2001). Partial match queries in relaxed multidimensional search trees. *Algorithmica* **29** 181–204.

[39] H. N. Minh and G. Jacob (2000). Symbolic integration of meromorphic differential systems via Dirichlet functions. *Discrete Mathematics* **210** 87–116.

[40] R. Neininger and L. Rüschendorf (1999). On the internal path length of d-dimensional quad trees. *Random Structures and Algorithms* **15** 25–41.

[41] R. Neininger and L. Rüschendorf (2001). Limit laws for partial match queries in quadtrees. *Annals of Applied Probability* **11** 452–469.

[42] V. V. Petrov (1975). *Sums of Independent Random Variables*. Springer, New York.

[43] H. Samet (1990a). *The Design and Analysis of Spatial Data Structures*. Addison-Wesley, Reading, MA.

[44] H. Samet (1990b). *Applications of Spatial Data Structures: Computer Graphics, Image Processing, and GIS*. Addison-Wesley, Reading, MA.

[45] R. Sedgewick (1980). *Quicksort*. Garland, New York.

Profiles of random trees: Limit theorems for random recursive trees and binary search trees

MICHAEL FUCHS[1]
Department of Applied Mathematics
National Chiao Tung University
Hsinchu 300
Taiwan

HSIEN-KUEI HWANG[2]
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

RALPH NEININGER[3]
Department of Mathematics
J.W. Goethe University
60325 Frankfurt a.M.
Germany

August 11, 2005

Abstract

We prove convergence in distribution for the profile (the number of nodes at each level), normalized by its mean, of random recursive trees when the limit ratio α of the level and the logarithm of tree size lies in $[0, e)$. Convergence of all moments is shown to hold only for $\alpha \in [0, 1]$ (with only convergence of finite moments when $\alpha \in (1, e)$). When the limit ratio is 0 or 1 for which the limit laws are both constant, we prove asymptotic normality for $\alpha = 0$ and a "quicksort type" limit law for $\alpha = 1$, the latter case having additionally a small range where there is no fixed limit law. Our tools are based on contraction method and method of moments. Similar phenomena also hold for other classes of trees; we apply our tools to binary search trees and give a complete characterization of the profile. The profiles of these random trees represent concrete examples for which the range of convergence in distribution differs from that of convergence of all moments.

1 Introduction

The profile or height profile of a tree is the sequence of numbers whose k-th element enumerates the number of nodes at distance k from the root of the tree (or the number of descendants in k-th generation in branching process terms). Profiles of trees are fine shape characteristics encountered in diverse problems

[1]Partially supported by National Science Council of ROC under the Grant `NSC-93-2119-M-009-003`.

[2]Partially supported by a research award of the Alexander von Humboldt Foundation and by National Science Council under the grant `NSC-92-2118-M-001-019`.

[3]Research supported by an Emmy Noether Fellowship of the DFG.

such as breadth-first search, data compression algorithms (Jacquet, Szpankowski, Tang, 2001), random generation of trees (Devroye and Robson, 1995), and the level-wise analysis of quicksort (Chern and Hwang, 2001b, Evans and Dunbar, 1982). In addition to their interest in applications and connections to many other shape parameters, we will show, through recursive trees and binary search trees, that profiles of random trees having roughly logarithmic height are a rich source of many intriguing phenomena. The high concentration of nodes at certain (log) levels results in the asymptotic bimodality for the variance, as already demonstrated in Drmota and Hwang (2005a); our purpose of this paper is to unveil and clarify the diverse phenomena exhibited by the limit distributions of the profiles of random recursive trees and binary search trees. The tools we use, as well as the results we derive, are of some generality.

Recursive trees. Recursive trees have been introduced as simple probability models for system generation (Na and Rapoport, 1970), spread of contamination of organisms (Meir and Moon, 1974), pyramid scheme (Bhattacharya and Gastwirth, 1984, Smythe and Mahmoud, 1995), stemma construction of philology (Najock and Heyde, 1982), Internet interface map (Janic et al., 2002), stochastic growth of networks (Chan et al., 2003). They are related to some Internet models (van Mieghem et al., 2001, van der Hofstad et al., 2001, Devroye, McDiarmid and Reed, 2002) and some physical models (Tetzlaff, 2002); they also appeared in Hopf algebra under the name of "heap-ordered trees"; see Grossman and Larson (1989). The bijection between recursive trees and binary search trees not only makes the former a flexible representation of the latter but also provides a rich direction for further extensions; see for example Mahmoud and Smythe (1995).

A simple way of constructing a random recursive tree of n nodes is as follows. One starts from a root node with the label 1; at stage i ($i = 2, \dots, n$) a new node with label i is attached uniformly at random to one of the previous nodes $(1, \dots, i-1)$. The process stops after node n is inserted. By construction, the labels of the nodes along any path from the root to a node form an increasing sequence; see Figure 2 for a recursive tree of 10 nodes. For a survey of probabilistic properties of recursive trees, see Smythe and Mahmoud (1995).

Known results for the profile of recursive trees. Let $X_{n,k}$ denote the number of nodes at level k in a random recursive tree of n nodes, where $X_{n,0} = 1$ (the root) for $n \ge 1$. Then $X_{n,k}$ satisfies (see van der Hofstad et al., 2002)

$$X_{n,k} \stackrel{\mathscr{D}}{=} X_{I_n,k-1} + X^*_{n-I_n,k}\ , \tag{1}$$

for $n, k \ge 1$ with $X_{n,0} = 1 - \delta_{n,0}$ ($\delta_{n,0}$ being Kronecker's symbol), where $(X_{n,k})$, $(X^*_{n,k})$ and (I_n) are independent, $X_{n,k} \stackrel{\mathscr{D}}{=} X^*_{n,k}$, and I_n is uniformly distributed over $\{1, \dots, n-1\}$.

Meir and Moon (1978) showed (implicitly) that

$$\mu_{n,k} := \mathbb{E}(X_{n,k}) = \frac{\mathsf{s}(n,k+1)}{(n-1)!} \qquad (0 \le k < n), \tag{2}$$

where $\mathsf{s}(n,k)$ denotes the unsigned Stirling numbers of the first kind; see also Moon (1974) and Dondajewski and Szymański (1982). By the approximations given in Hwang (1995), we then have

$$\mu_{n,k} = \frac{\lambda_n^k}{\Gamma(1+\alpha_{n,k})k!}\left(1 + O\left(\lambda_n^{-1}\right)\right), \tag{3}$$

uniformly for $1 \le k \le K\lambda_n$, for any $K > 1$, where, *here and throughout this paper,*

$$\lambda_n := \max\{\log n, 1\}, \qquad \alpha_{n,k} := k/\lambda_n,$$

and Γ denotes the Gamma function. This approximation implies, in particular, a local limit theorem for the depth (distance of a random node to the root); see Devroye (1998), Szymański (1990), Mahmoud (1991).

The second moment is also implicit in Meir and Moon (1978)

$$\mathbb{E}(X_{n,k}^2) = \sum_{0 \le j \le k} \binom{2j}{j} \frac{\mathsf{s}(n, k+j+1)}{(n-1)!};$$

see also van der Hofstad et al. (2002). Precise asymptotic approximations for the variance $\mathbb{V}(X_{n,k})$ were derived in Drmota and Hwang (2005a) for all ranges of k. In particular, the variance is asymptotically of the same order as $\mu_{n,k}^2$ when $\alpha \in (0,2)$ except $k \sim \lambda_n$ (where the profile variance exhibits a bimodal behavior).

Limit distribution when $0 \le \alpha < e$**.** From the asymptotic estimate (3), we have

$$\frac{\log \mu_{n,k}}{\lambda_n} \to \alpha - \alpha \log \alpha,$$

where *here and throughout this paper* $k = k(n)$ and $\alpha := \lim_{n\to\infty} k(n)/\lambda_n$. Thus $\mu_{n,k} \to \infty$ when $\alpha < e$. Note that the expected height (length of the longest path from the root) of random recursive trees is asymptotic to $e\lambda_n$; see Devroye (1987) or Pittel (1994).

Define a class of random variables $X(\alpha)$ by the fixed-point equation

$$X(\alpha) \stackrel{\mathscr{D}}{=} \alpha U^\alpha X(\alpha) + (1-U)^\alpha X(\alpha)^*, \tag{4}$$

with $\mathbb{E}(X(\alpha)) = 1$, where $X(\alpha), X(\alpha)^*, U$ are independent, $X(\alpha)^* \stackrel{\mathscr{D}}{=} X(\alpha)$, and U is uniformly distributed in the unit interval; see Proposition 1 for existence and properties of $X(\alpha)$. Define $X(0) = 1$.

Theorem 1. *(i) If $0 \le \alpha < e$, then*

$$\frac{X_{n,k}}{\mu_{n,k}} \xrightarrow{\mathscr{D}} X(\alpha), \tag{5}$$

where $\xrightarrow{\mathscr{D}}$ denotes convergence in distribution.

(ii) If $0 \le \alpha < m^{1/(m-1)}$, where $m \ge 2$, then $X_{n,k}/\mu_{n,k}$ converges to $X(\alpha)$ with convergence of the first m moments but not the $(m+1)$-st moment.

In particular, convergence of the second moment holds for $0 \le \alpha < 2$.

Corollary 1. *If $0 \le \alpha < 2$, then*

$$\mathbb{V}(X_{n,k}) \sim \left(\frac{\Gamma(\alpha+1)^2}{(1-\alpha/2)\Gamma(2\alpha+1)} - 1 \right) \mu_{n,k}^2.$$

Note that the coefficient on the right-hand side becomes zero when $\alpha = 0$ and $\alpha = 1$, and the variance indeed exhibits a *bimodal behavior* when $\alpha = 1$; see Figure 1 for a plot and Drmota and Hwang (2005a) or below for more precise approximations to the variance.

Since $m^{1/(m-1)} \downarrow 1$, the unit interval is the only range where convergence of all moments holds.

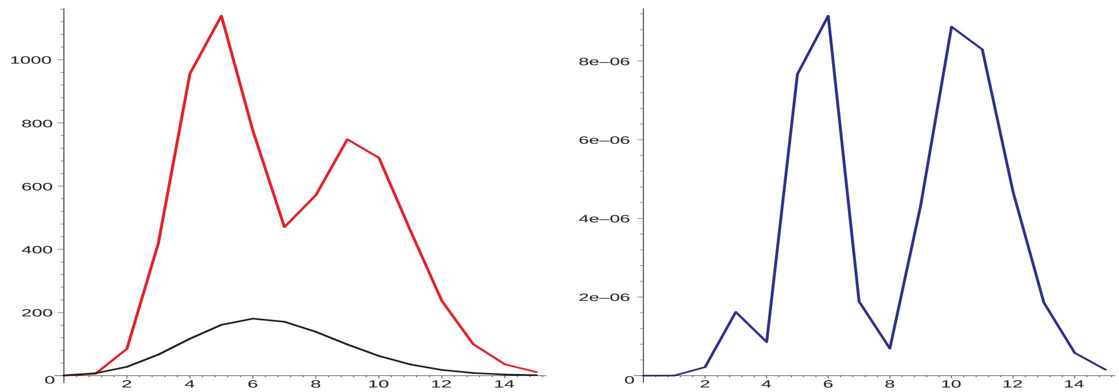

Figure 1: *A plot of* $\mathbb{E}(X_{n,k})$ *(the unimodal curve),* $\mathbb{V}(X_{n,k})$ *(the bimodal curve with higher valley), and* $|\mathbb{E}(X_{n,k} - \mu_{n,k})^3|$ *(right) of the number* $X_{n,k}$ *of nodes at level* k *in random recursive trees of* $n = 1100$ *nodes, all normalized by their maximum values. Note that the valley of* $|\mathbb{E}(X_{1100,k} - \mu_{1100,k})^3|$ *(when normalized by* n^3*) is deeper than that of* $\mathbb{V}(X_{1100,k})$ *(normalized by* n^2*); see Corollary 5 for the general description.*

Corollary 2. *If* $0 \le \alpha \le 1$*, then*

$$\frac{X_{n,k}}{\mu_{n,k}} \xrightarrow{\mathscr{M}} X(\alpha), \tag{6}$$

where $\xrightarrow{\mathscr{M}}$ *denotes convergence of all moments. Convergence of all moments fails for* $1 < \alpha < e$.

Thus the profile of random recursive trees represents a concrete example for which *the range of convergence in distribution is different from that of convergence of all moments.* We will show that such a property also holds for random binary search trees; it is expected to hold for other trees like ordered (or plane) recursive trees and m-ary search trees, but the technicalities are expected to be much more complicated. We focus at this stage on new phenomena and their proofs, not on generality.

The proof of (5) relies on the contraction method developed in Neininger and Rüschendorf (2004) (see also the survey paper Rösler and Rüschendorf, 2001), and the moment convergence $X_{n,k}/\mu_{n,k}$ uses the method of moments. Both methods are technically more involved because we are dealing with recurrences with two parameters. We will indeed prove a stronger approximation to (5) by deriving a rate under the Zolotarev metric (see Zolotarev, 1976).

But why $m^{1/(m-1)}$? This is readily seen by the recurrence of the moments $\nu_m(\alpha) := \mathbb{E}(X(\alpha)^m)$ of $X(\alpha)$

$$\nu_m(\alpha) = \frac{1}{m - \alpha^{m-1}} \sum_{1 \le h < m} \binom{m}{h} \nu_h(\alpha)\nu_{m-h}(\alpha)\alpha^{h-1} \frac{\Gamma(h\alpha + 1)\Gamma((m-h)\alpha + 1)}{\Gamma(m\alpha + 1)} \qquad (m \ge 2), \tag{7}$$

where $\nu_0(\alpha) = \nu_1(\alpha) = 1$. This recurrence is well-defined for $\nu_m(\alpha)$ when $\alpha < m^{1/(m-1)}$. This explains the special sequence $m^{1/(m-1)}$.

Note that since $\mathbb{E}(X(\alpha)^m) = \infty$ for $\alpha \ge m^{1/(m-1)}$, we have $\mathbb{E}(X_{n,k}/\mu_{n,k})^m \to \infty$ in that range.

A "quicksort-type" limit distribution when $\alpha = 1$**.** Since $X(1) = 1$, we can refine the limit result (5) for $\alpha = 1$ as follows.

Theorem 2. *(i) If $k = \lambda_n + t_{n,k}$, where $|t_{n,k}| \to \infty$ and $t_{n,k} = o(\lambda_n)$, then*

$$\frac{X_{n,k} - \mu_{n,k}}{t_{n,k}\lambda_n^{k-1}/k!} \xrightarrow{\mathscr{M}} X'(1), \tag{8}$$

where $X'(1) := (d/d\alpha)X(\alpha)|_{\alpha=1}$ satisfies

$$X'(1) \stackrel{\mathscr{D}}{=} UX'(1) + (1-U)X'(1)^* + U + U\log U + (1-U)\log(1-U),$$

with $X'(1), X'(1)^*, U$ independent and $X'(1) \stackrel{\mathscr{D}}{=} X'(1)^*$.

(ii) If $k = \lambda_n + O(1)$, then the sequence of random variables $(X_{n,k} - \mu_{n,k})/\sqrt{\mathbb{V}(X_{n,k})}$ does not converge to a fixed law.

Although (8) can also be proved by the contraction method, we prove both results of the theorem by the method of moments because the proof for the non-convergence part is readily modified from that for (8); see also Chern et al. (2002) for more examples having no convergence to fixed limit law. On the other hand, since the distribution of $X'(1)$ is uniquely characterized by its moment sequence (see (41)), we have the convergence in distribution as follows.

Corollary 3. *If $k = \lambda_n + t_{n,k}$, where $|t_{n,k}| \to \infty$ and $t_{n,k} = o(\lambda_n)$, then*

$$\frac{X_{n,k} - \mu_{n,k}}{t_{n,k}\lambda_n^{k-1}/k!} \xrightarrow{\mathscr{D}} X'(1).$$

The same limit law $X'(1)$ also appeared in the total path length (which is $\sum_k kX_{n,k}$) of recursive trees (see Dobrow and Fill, 1999), or essentially the total left path length of random binary search trees, and the cost of an in-situ permutation algorithm; see Hwang and Neininger (2002).

The appearance of the same limit law as the total path length is not a coincidence. *Intuitively*, almost all nodes lie at the levels $k = \lambda_n + O(\sqrt{\lambda_n})$ (since $\mathbb{E}(X_{n,k}) \asymp n/\sqrt{\lambda_n}$ by (3)) and it is these nodes that contribute predominantly to the total path length; see also (9) below for an estimate of the variance. *Analytically*, a deeper connection between the profile and the total path length is seen through the level polynomials $\sum_k X_{n,k}z^k$ (properly normalized) for which we can derive, following Chauvin et al. (2001), an almost sure convergence to some (complex-valued) limit random variable. From such a uniform convergence, the profile is quickly linked to the total path length by taking derivative of the normalized level polynomial with respect to z and substituting $z = 1$. Indeed, limit theorems for weighted path-lengths of the form $\sum_k k^m X_{n,k}$, as well as the width ($\max_k X_{n,k}$), can be obtained as by-products. These and finer results on correlations and expected width are discussed in Drmota and Hwang (2005b).

Asymptotics of the variance. As a consequence of our convergence of all moments, we have the following estimate for the variance.

Corollary 4. *If $k = \lambda_n + t_{n,k}$, where $t_{n,k} = o(\lambda_n)$, then the variance of $X_{n,k}$ satisfies*

$$\mathbb{V}(X_{n,k}) \sim p_2(t_{n,k})\left(\frac{\lambda_n^{k-1}}{k!}\right)^2, \tag{9}$$

where $p_2(t_{n,k}) := c_2t_{n,k}^2 + 2c_1t_{n,k} + c_0$ with

$$c_2 := 2 - \frac{\pi^2}{6}, \quad c_1 := c_2(1-\gamma) - \zeta(3) + 1$$

$$c_0 := c_2\left(\gamma^2 - 2\gamma + 3\right) - 2(\zeta(3) - 1)(1-\gamma) - \frac{\pi^4}{360}. \tag{10}$$

Here γ denotes Euler's constant and $\zeta(3) := \sum_{j\ge1} j^{-3}$.

The expression (9) explains the valley for the variance in Figure 1. Note that $\mathbb{V}(X_{n,k})/\mu_{n,k}^2 = O(t_{n,k}^2/\lambda_n^2)$ when $t_{n,k} = o(\lambda_n)$.

Our proof indeed yields the following extremal orders of $|\mathbb{E}(X_{n,k} - \mu_{n,k})^m|$ for $m \geq 2$.

Corollary 5. *The absolute value of the m-th central moment satisfies*

$$\max_{0 \leq k < n} |\mathbb{E}(X_{n,k} - \mu_{n,k})^m| \asymp \lambda_n^{-m} n^m,$$
$$\min_{|k-\lambda_n| = O(\sqrt{\lambda_n})} |\mathbb{E}(X_{n,k} - \mu_{n,k})^m| \asymp \lambda_n^{-3m/2} n^m,$$

where the maximum is achieved at $k = \lambda_n \pm \sqrt{\lambda_n}(1 + o(1))$ *and the minimum at* $k = \lambda_n + O(1)$.

More refined results can be derived as in Drmota and Hwang (2005a). For example, by (40) below, we have

$$\max_{0 \leq k < n} |\mathbb{E}(X_{n,k} - \mu_{n,k})^m| \sim |\mathbb{E}(X'(1)^m)| e^{-m/2} \left(\frac{n}{\sqrt{2\pi}\lambda_n} \right)^m,$$

for $m \geq 2$, where $\mathbb{E}(X'(1)^m)$ can be computed recursively; see (41).

Asymptotic normality when $\alpha = 0$**.** The profile $X_{n,k}$ in the remaining range $1 \leq k = o(\lambda_n)$ will be shown to be asymptotically normally distributed. It is known (see Bergeron et al., 1992) that the out-degree of the root $X_{n,1}$ satisfies

$$\mathbb{P}(X_{n,1} = j) = \frac{\mathsf{s}(n-1, j)}{(n-1)!} \qquad (1 \leq j < n);$$

thus $X_{n,1}$ is asymptotically normal with mean and variance both asymptotic to λ_n. Equivalently, $X_{n,1}$ is the number of nodes on the rightmost branch (the path starting from the root and always going right until reaching an external node) in a random binary search trees of $n - 1$ nodes; see the transformation below for more information.

Let $\Phi(x) := (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2} \,\mathrm{d}t$ denote the distribution function of the standard normal distribution.

Theorem 3. *The distribution of the profile* $X_{n,k}$ *satisfies*

$$\sup_x \left| \mathbb{P}\left(\frac{X_{n,k} - \lambda_n^k/k!}{\lambda_n^{k-1/2}/\sqrt{(k-1)!^2(2k-1)}} < x \right) - \Phi(x) \right| = O\left(\sqrt{\frac{k}{\lambda_n}} \right), \tag{11}$$

uniformly for $1 \leq k = o(\lambda_n)$*, with mean and variance asymptotic to*

$$\begin{cases} \mathbb{E}(X_{n,k}) \sim \dfrac{\lambda_n^k}{k!}, \\ \mathbb{V}(X_{n,k}) \sim \dfrac{\lambda_n^{2k-1}}{(k-1)!^2(2k-1)}. \end{cases}$$

In particular, $X_{n,2}$ is asymptotically normally distributed with mean asymptotic to $\frac{1}{2}\lambda_n^2$ and variance to $\frac{1}{3}\lambda_n^3$. A similar central limit theorem appeared in the logarithmic order of a random element in symmetric groups; see Erdős and Turán (1967).

Unlike previous cases, the proof of this result is based on a polynomial decomposition of the associated generating functions using characteristic functions and singularity analysis (see Flajolet and Odlyzko,

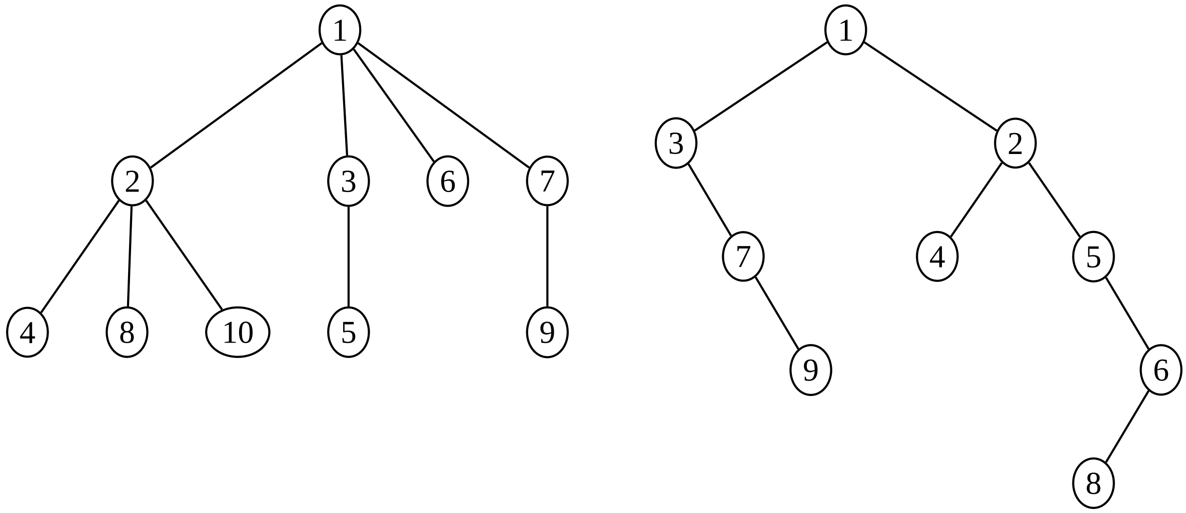

Figure 2: *A recursive tree of* 10 *nodes and its corresponding transformed binary increasing tree of* 9 *nodes.*

1990), the reasons being (i) this method leads to the optimal Berry-Esseen bound (11), which is not obvious by the method of moments; (*ii*) it is of independent methodological interests, and (*iii*) it can also be applied to give an alternative proof of (6).

The asymptotic normality of $X_{n,k}$ when $\alpha = 0$ indicates that nodes are generated in a very regular way in recursive trees, at least for the first $o(\lambda_n)$ levels. The rough picture here is that each node at these levels "attracts" about λ_n/k new-coming nodes, as is obvious from (3); see also Drmota and Hwang (2005b) for an asymptotic independence property for the number of nodes at two different levels, both being $o(\lambda_n)$ away from the root.

Profiles of random binary search trees. Binary search trees are one of the most studied fundamental data structures in Computer Algorithms. They have also been introduced in other fields under different forms; see Drmota and Hwang (2005a) for more references.

This tree model is characterized by a recursive splitting process in which $n \geq 2$ distinct labels are split into a root and two subtrees formed recursively by the same procedure (one may be empty) of sizes J_n and $n-1-J_n$, where J_n is uniformly distributed in $\{0, 1, \ldots, n-1\}$. Such a model is isomorphic to *binary increasing trees* in which a sequence of $n \geq 2$ continuous random variables (independent and identically distributed) is split into a root with the smallest label and two subtrees formed recursively by the same splitting process corresponding to the subsequences to the left and right respectively of the smallest label. Note that when given a random permutation of n elements the size of the left subtree of the binary increasing tree constructed from the permutation equals j, $0 \leq j \leq n-1$ with equal probability $1/n$, the same as in random binary search trees.

A recursive tree can be transformed into a binary increasing tree by the well-known procedure (referred to as the *natural correspondence* in Kunth, 1997 and the *rotation correspondence* by others): drop first the root and arrange all subtrees from left to right in increasing order of their root labels; sibling relations are transformed into right branches (of the leftmost node in that generation) and the leftmost branches remain unchanged; a final relabeling (using labels from 1 to $n-1$) of nodes then yields a binary increasing tree of $n-1$ nodes. Such a transformation is invertible; see Figure 2.

Under this transformation, the profile $X_{n,k}$ in recursive trees becomes essentially the number of nodes in random binary search trees of $n-1$ nodes with left-distance $k-1$ ($k \geq 1$), the *left-distance* of a node being the number of left-branches needed to traverse from the root to that node. This also explains the recurrence (1).

Known and new results for profiles of random binary search trees. We distinguish two types of nodes for binary search trees: external nodes $Y_{n,k}$ (virtual nodes completed so that all nodes are of out-degree either zero or two) and internal nodes $Z_{n,k}$ (nodes holding labels). Chauvin et al. (2001) established *almost sure convergence* for $Y_{n,k}/\mathbb{E}(Y_{n,k})$ and $Z_{n,k}/\mathbb{E}(Z_{n,k})$ when $1.2 \le \alpha \le 2.8$, and recently Chauvin et al. (2005) extended the range for $Y_{n,k}/\mathbb{E}(Y_{n,k})$ to the optimal range $\alpha_- < \alpha < \alpha_+$, the two numbers $\alpha_- \approx 0.37, \alpha_+ \approx 4.31$ being the fill-up and height constants (of binary search trees), namely, $0 < \alpha_- < 1 < \alpha_+$ solving the equation $e^{(z-1)/z} = z/2$; see also Chauvin and Rouault (2004). For other known results on the profiles $Y_{n,k}$, see Drmota and Hwang (2005a) and the references therein.

Our tools for recursive trees also apply to binary search trees. Briefly, we derive convergence in distribution for $Y_{n,k}/\mathbb{E}(Y_{n,k})$ and $Z_{n,k}/\mathbb{E}(Z_{n,k})$ in the range $\alpha \in (\alpha_-, \alpha_+)$ and convergence of all moments for $\alpha \in [1, 2]$, the degenerate cases $\alpha = 1, 2$ being further refined by more explicit limit laws; see Section 7 for details.

While it is expected that the profiles for both types of nodes have similar behaviors to $X_{n,k}$, we will derive finer results showing more delicate structural difference between internal nodes and external nodes.

Organization of the paper. Since most of our asymptotic approximations are based on the solution (exact or asymptotic) of the underlying double-indexed recurrence (in n and k), we start from solving the recurrence in the next section. The proof of the convergence in distribution (5) of $X_{n,k}/\mu_{n,k}$ when $0 < \alpha < e$ by contraction method is given in Section 3. Then we prove the moment convergence part of Theorem 1 in Section 4 and Theorem 2 in Section 5. The asymptotic normality when $\alpha = 0$ is proved in Section 6, where an alternative proof of (6) is also indicated. Our methods of proof can be easily amended for binary search trees, and the results are given in Section 7. We conclude this paper with a few questions.

Notations. Throughout this paper, $\lambda_n := \max\{\log n, 1\}$, $\alpha_{n,k} := k/\lambda_n$ and $\alpha := \lim_{n\to\infty} \alpha_{n,k}$ when the limit exists. The symbol $[z^n]f(z)$ stands for the coefficient of z^n in the Taylor expansion of $f(z)$. The generic symbols ε and K always represent sufficiently small and large, respectively, positive constants whose values may vary from one occurrence to another. Finally, U represents a uniform $[0, 1]$ random variable.

2 The double-indexed recurrence and asymptotic transfer

Since all moments (centered or not) satisfy the same recurrence, we derive in this section the exact solution and study a simple type of asymptotic transfer (relating the asymptotics of the recurrence to that of the non-homogeneous part) for such a recurrence.

By (1), we have the recurrence for the probability generating functions $P_{n,k}(y) := \mathbb{E}(y^{X_{n,k}})$

$$P_{n,k}(y) = \frac{1}{n-1} \sum_{1 \le j < n} P_{j,k-1}(y) P_{n-j,k}(y) \qquad (n \ge 2; k \ge 1), \tag{12}$$

with $P_{n,0}(y) = y$ for $n \ge 1$ and $P_{0,k}(y) = 1$.

Recurrence of factorial moments. Let

$$A_{n,k}^{(m)} := \mathbb{E}(X_{n,k}(X_{n,k}-1)\cdots(X_{n,k}-m+1)) = P_{n,k}^{(m)}(1).$$

Then $A_{n,k}^{(0)} = 1$ for $n,k \geq 0$. By (12), we have the recurrence

$$A_{n,k}^{(m)} = \frac{1}{n-1}\sum_{1\leq j<n}\left(A_{j,k-1}^{(m)} + A_{j,k}^{(m)}\right) + B_{n,k}^{(m)} \qquad (n\geq 2; k,m\geq 1),$$

where

$$B_{n,k}^{(m)} = \sum_{1\leq h<m}\binom{m}{h}\frac{1}{n-1}\sum_{1\leq j<n} A_{j,k-1}^{(h)}A_{n-j,k}^{(m-h)}, \tag{13}$$

with the boundary conditions $A_{n,0}^{(1)} = 1$ for $n \geq 1$ and $A_{n,0}^{(m)}(0) = 0$ for $m \geq 2$ and $n \geq 1$.

Exact solution of the recurrence. Consider a recurrence of the form

$$a_{n,k} = \frac{1}{n-1}\sum_{1\leq j<n}\left(a_{j,k} + a_{j,k-1}\right) + b_{n,k}, \qquad (n\geq 2; k\geq 1), \tag{14}$$

with $a_{1,k}$ and $b_{n,k}$ given. We assume, without loss of generality, that $a_{0,k} = 0$ (otherwise, we need only to modify the values of $a_{1,k}$ and $b_{n,k}$).

Lemma 1. *For $n \geq 1$ and $k \geq 0$,*

$$a_{n,k} = b_{n,k} + \sum_{1\leq j<n}\sum_{0\leq r\leq k}\frac{b_{j,k-r}}{j}[u^r](u+1)\prod_{j<\ell<n}\left(1+\frac{u}{\ell}\right), \tag{15}$$

where $b_{1,k} := a_{1,k}$.

Proof. Let $a_n(u) := \sum_k a_{n+1,k}u^k$ and $b_n(u) := \sum_k b_{n+1,k}u^k$. Then $a_n(u)$ satisfies the recurrence

$$a_n(u) = \frac{1+u}{n}\sum_{0\leq j<n} a_j(u) + b_n(u) \qquad (n\geq 1),$$

with the initial condition $a_0(u) = \sum_k a_{1,k}u^k$. By taking the difference $na_n(u) - (n-1)a_{n-1}(u)$, we obtain

$$a_n(u) = \left(1+\frac{u}{n}\right)a_{n-1}(u) + b_n(u) - \frac{n-1}{n}b_{n-1}(u) \qquad (n\geq 2).$$

Solving this linear recurrence yields

$$a_n(u) = b_n(u) + (1+u)\sum_{0\leq j<n}\frac{b_j(u)}{j+1}\prod_{j+2\leq\ell\leq n}\left(1+\frac{u}{\ell}\right) \qquad (n\geq 1),$$

(since $b_0(u) := a_0(u)$). Taking coefficient of u^k on both sides leads to (15). ∎

Mean value. Applying (15) with $b_{n,k} = \delta_{n,1}\delta_{0,k}$, we obtain for $n \geq 1$ and $k \geq 0$

$$\begin{aligned}\mu_{n,k} &= [u^k] \prod_{1\leq \ell < n} \left(1 + \frac{u}{\ell}\right) \\ &= \frac{\mathsf{s}(n, k+1)}{(n-1)!}.\end{aligned} \tag{16}$$

This rederives (2).

A uniform estimate for the expected profile. For later use, we derive a uniform bound for $\mu_{n,k}$.

Lemma 2. *The mean satisfies*

$$\mu_{n,k} = O\left((v\lambda_n)^{-1/2} v^{-k} n^v\right), \tag{17}$$

uniformly for $1 \leq k < n$, *where* $0 < v = O(1)$.

Proof. Note that by (16), we have the obvious inequality

$$\mu_{n,k} v^k \leq \prod_{1\leq \ell < n} \left(1 + \frac{v}{\ell}\right) \qquad (v > 0),$$

which leads to $\mu_{n,k} = O\left(v^{-k} n^v\right)$ for $1 \leq k < n$. But this is too crude for our purpose.

By Cauchy's integral formula,

$$\begin{aligned}\mu_{n,k} &\leq \frac{v^{-k}}{2\pi} \int_{-\pi}^{\pi} \prod_{1\leq \ell \leq n} \left|1 + \frac{ve^{it}}{\ell}\right| \mathrm{d}t \\ &\leq \frac{v^{-k}}{2\pi} \int_{-\pi}^{\pi} \exp\left(v(\cos t) \sum_{1\leq \ell \leq n} \frac{1}{\ell} + O(1)\right) \mathrm{d}t \\ &= O\left((v\lambda_n)^{-1/2} v^{-k} n^v\right).\end{aligned}$$

proving (17). ∎

Note that when $k = O(\lambda_n)$, then the right-hand side of (17) is optimal if we take $v = k/\lambda_n$ and (17) becomes $\mu_{n,k} = O(\lambda_n^k/k!)$. Thus (17) is tight when $k = O(\lambda_n)$. This also explains why we write $(v\lambda_n)^{-1/2}$ instead of $\lambda_n^{-1/2}$ (to keep uniformity when $k = o(\lambda_n)$ and we choose $v = k/\lambda_n$).

On the other hand, leaving v unspecified in (17) and in many other estimates in this paper considerably simplifies the analysis.

A simple asymptotic transfer. We will need the following result when applying the contraction method. It roughly says that when the non-homogeneous part $b_{n,k}$ of (14) is of order $\mu_{n,k}^w$, where $w > 1$, then $a_{n,k}$ is also of the same order for certain range of α.

Lemma 3. *If* $b_{n,k} = O\left(((v\lambda_n)^{-1/2} v^{-k} n^v)^w\right)$ *for all* $1 \leq k \leq n$, *where* $w > 1$ *and* $0 < v < v_0$, *then*

$$a_{n,k} = O\left(\frac{1}{w - v^{w-1}} \left((v\lambda_n)^{-1/2} v^{-k} n^v\right)^w\right),$$

uniformly for $1 \leq k \leq n$, *provided that* $0 < v < \min\{w^{1/(w-1)}, v_0\}$. *Similarly, replacing* O *by* o *in the estimate for* $b_{n,k}$ *yields an* o*-estimate for* $a_{n,k}$.

Proof. By the exact expression for $a_{n,k}$, we have, for $0 < v < v_0$,

$$a_{n,k} - b_{n,k} = O\left(\sum_{1\le j<n}\sum_{0\le r\le k}\frac{1}{j}\left((v\lambda_j)^{-1/2}v^{-k+r}j^{v}\right)^{w}[u^r](1+u)\prod_{j<\ell<n}\left(1+\frac{u}{\ell}\right)\right). \tag{18}$$

The inner sum over r can be simplified as follows.

$$\begin{aligned}\sum_{0\le r\le k} v^{-(k-r)w}[u^r](1+u)\prod_{j<\ell<n}\left(1+\frac{u}{\ell}\right) &\le v^{-kw}\sum_{r\ge 0}v^{rw}[u^r](1+u)\prod_{j<\ell<n}\left(1+\frac{v^w t}{\ell}\right)\\ &= v^{-kw}(1+v^w)\prod_{j<\ell<n}\left(1+\frac{v^w}{\ell}\right)\\ &= O\left(v^{-kw}\left(\frac{n}{j}\right)^{v^w}\right),\end{aligned} \tag{19}$$

uniformly in j. Substituting this estimate into (18), we obtain

$$\begin{aligned}a_{n,k} &= O\left(\left((v\lambda_n)^{-1/2}v^{-k}n^{v}\right)^{w} + v^{-kw}n^{v^w}\sum_{1\le j<n}(v\lambda_j)^{-w/2}j^{wv-v^w-1}\right)\\ &= O\left(\frac{1}{w-v^{w-1}}\left((v\lambda_n)^{-1/2}v^{-k}n^{v}\right)^{w}\right),\end{aligned}$$

uniformly for $1 \le k \le n$, where $0 < v < w^{1/(w-1)}$. The o-estimate is similarly proved. This completes the proof of Lemma 3. ∎

3 Convergence in distribution when $0 < \alpha < e$

We prove the first part of Theorem 1 (excepting $\alpha = 0$) in this section by contraction method based on the framework developed in Neininger and Rüschendorf (2004). The new difficulty arising here is the asymptotics of the double-indexed recurrence (14) (instead of single-indexed ones previously encountered).

The underlying idea. The idea used here is roughly as follows.

Define $\bar{X}_{n,k} := X_{n,k}/\mu_{n,k}$. Then, by (1), $\bar{X}_{n,k}$ satisfies the recurrence

$$\bar{X}_{n,k} \stackrel{\mathscr{D}}{=} \frac{\mu_{I_n,k-1}}{\mu_{n,k}}\bar{X}_{I_n,k-1} + \frac{\mu_{n-I_n,k}}{\mu_{n,k}}\bar{X}^*_{n-I_n,k}, \tag{20}$$

with independence conditions as in (1). By the estimates (3) and the relation $I_n = \lceil (n-1)U\rceil$, we expect that

$$\frac{\mu_{I_n,k-1}}{\mu_{n,k}} \approx \frac{k}{\lambda_n}\left(\frac{\lambda_n+\log U}{\lambda_n}\right)^{k-1} \to \alpha U^{\alpha},$$

with suitable meaning for the convergence; similarly,

$$\frac{\mu_{n-I_n,k}}{\mu_{n,k}} \to (1-U)^{\alpha}.$$

Thus if we expect that $\bar{X}_{n,k} \to X(\alpha)$, then $X(\alpha)$ satisfies the fixed-point equation (4).

To justify these steps, we apply the contraction method.

Contraction method. The fixed-point equation (4) has a few special properties not enjoyed by single-indexed recursions encountered in the literature for which the typical fixed-point equation has the form

$$X \stackrel{\mathscr{D}}{=} \sum_{1 \le j \le h} C_j X^{(j)} + b, \tag{21}$$

with $X^{(1)}, \ldots, X^{(h)}, (C_1, \ldots, C_h, b)$ independent, $X^{(j)} \stackrel{\mathscr{D}}{=} X$, and $0 \le C_j \le 1$ almost surely for all $1 \le j \le h$. Here, h may be deterministic or integer-valued random variables. The special range $[0, 1]$ for the coefficients $C_1 \ldots, C_j$ is roughly due to the relation

$$\frac{\sigma(I_j^{(n)})}{\sigma(n)} \to C_j,$$

where, in various applications (see Neininger and Rüschendorf, 2004), σ is the leading term in the expansion of the standard deviation of the underlying random variable and $0 \le I_j^{(n)} \le n$ are the sizes of the subproblems. Typically, σ is a monotonically increasing function, hence we obtain $0 \le C_j \le 1$.

In general, the Lipschitz constant of the map of probability measures associated with (21) under the Zolotarev metric ζ_w is assessed by $\sum_j \mathbb{E}(C_j^w)$. This term is monotonically decreasing as w increases. Thus, in typical applications for which one expects a contraction, the sum $\sum_j \mathbb{E}(C_j^w)$ has to satisfy $\sum_j \mathbb{E}(C_j^w) < 1$, and for that purpose, one has to choose w sufficiently large; see Neininger and Rüschendorf (2004) for implications of this condition on the moments required.

For the bi-indexed recursion of $X_{n,k}$, we are led to the fixed-point equation (4), where the coefficient αU^α may have values larger than one for $\alpha > 1$. This implies that the corresponding estimate $\mathbb{E}(\alpha U)^w + \mathbb{E}(1-U)^w$ for the Lipschitz constant is not decreasing in w. When $\alpha < e$ increases, the range where we have contraction becomes smaller and vanishes in the boundary case $\alpha = e$.

Notations. We denote by $\mathcal{M}$ the space of univariate probability measures, by $\mathcal{M}_w \subset \mathcal{M}$ the space of probability measures with finite absolute w-th moment, and by $\mathcal{M}_w(1) \subset \mathcal{M}_w$ the subspace of probability measures with unit mean, where $1 < w \le 2$. Zolotarev [50] introduced a family of metrics ζ_w, which, for $1 < w \le 2$ are given by

$$\zeta_w(\nu_1, \nu_2) = \sup_{f \in \mathcal{F}_w} |\mathbb{E}(f(X) - f(Y))|, \qquad (\nu_1, \nu_2 \in \mathcal{M}_w(1)),$$

where X and Y have the distributions $\mathcal{L}(X) = \nu_1$, $\mathcal{L}(Y) = \nu_2$.

We have

$$\mathcal{F}_w := \{f \in C^1(\mathbb{R}, \mathbb{R}) \,:\, |f'(x) - f'(y)| \le |x-y|^{w-1}\},$$

with $C^1(\mathbb{R}, \mathbb{R})$ the space of continuously differentiable functions on $\mathbb{R}$. We will use the property that convergence in ζ_w implies weak convergence and that ζ_w is ideal of order w, i.e., we have for W independent of (X, Y) and $c \ne 0$

$$\zeta_w(X + W, Y + W) \le \zeta_w(X, Y), \qquad \zeta_w(cX, cY) = |c|^w \zeta_w(X, Y).$$

For general reference and properties of ζ_w, see Zolotarev [51] and Rachev [43].

We also use the minimal L_p metrics ℓ_p, defined for $1 < p \le 2$ by

$$\ell_p(\nu_1, \nu_2) = \inf\{\|X - Y\|_p \,:\, \mathcal{L}(X) = \nu_1, \mathcal{L}(Y) = \nu_2\}, \qquad (\nu_1, \nu_2 \in \mathcal{M}_p),$$

where $\|X\|_p$ denotes the L_p-norm of a random variable X. For simplicity, we use the abbreviation $\zeta_w(X, Y) := \zeta_w(\mathcal{L}(X), \mathcal{L}(Y))$ for ζ_w as well as for the other metrics appearing subsequently.

In addition, we assume that

$$R(n) := |k - \alpha\lambda_n| = |\alpha_{n,k} - \alpha|\lambda_n = o(\lambda_n),$$

where $0 < \alpha < e$, and fix a constant s as follows. If $2 \le \alpha < e$, then $1 < s < \rho$ with $\rho \in (1, 2]$ the unique solution of $\rho = \alpha^{\rho-1}$, and $s := 2$ if $0 < \alpha < 2$. The bound ρ also identifies the best possible order for the existence of absolute moment of $X(\alpha)$. Note that s satisfies $s - \alpha^{s-1} > 0$, which is the continuous version of $m - \alpha^{m-1} > 0$ appearing in (7).

Properties of $X(\alpha)$. Define the map

$$T : \mathcal{M} \to \mathcal{M}, \quad \nu \mapsto \mathcal{L}(\alpha U^\alpha Z + (1-U)^\alpha Z^*),$$

where Z, Z^*, U are independent, $\mathcal{L}(Z) = \mathcal{L}(Z^*) = \nu$.

Proposition 1. *For $0 < \alpha < e$, the restriction of T to $\mathcal{M}_s(1)$ has a unique fixed point $\mathcal{L}(X(\alpha))$. Furthermore, $\mathbb{E}|X(\alpha)|^\rho = \infty$ for $2 \le \alpha < e$.*

Proof. By Lemma 3.1 in Neininger and Rüschendorf (2004), T is a Lipschitz map in ζ_s with Lipschitz constant bounded above by

$$\operatorname{lip}(T) \le \frac{\alpha^s + 1}{\alpha s + 1}.$$

Thus $\operatorname{lip}(T) < 1$ by our choice of s. Also T has a unique fixed point in the subspace $\mathcal{M}_s(1)$ by Lemma 3.3 in Neininger and Rüschendorf (2004).

When $2 \le \alpha < e$, we assume $\mathbb{E}|X(\alpha)|^\rho < \infty$ and prove a contradiction. First we have $\mathbb{E}|X(\alpha)|^\rho = \mathbb{E}|\alpha U^\alpha X(\alpha) + (1-U)^\alpha X(\alpha)^*|^\rho$, where $X(\alpha), X(\alpha)^*, U$ are independent with $\mathcal{L}(X(\alpha)) = \mathcal{L}(X(\alpha)^*)$. Note that $X(\alpha) \ge 0$ almost surely. Furthermore, $\mathbb{E}(X(\alpha)) = 1$ implies that there is a set with positive probability in which we have $X(\alpha) > 0$ and $X(\alpha)^* > 0$. It follows that

$$\begin{aligned}\mathbb{E}|X(\alpha)|^\rho = \mathbb{E}(X(\alpha)^\rho) &= \mathbb{E}(\alpha U^\alpha X(\alpha) + (1-U)^\alpha X(\alpha)^*)^\rho \\ &> \mathbb{E}\left(\alpha^\rho U^{\alpha\rho} X(\alpha)^\rho + (1-U)^{\alpha\rho}(X(\alpha)^*)^\rho\right) \\ &= \frac{\alpha^\rho + 1}{\alpha\rho + 1}\mathbb{E}(X(\alpha)^\rho) \\ &= \mathbb{E}(X(\alpha)^\rho),\end{aligned}$$

by the definition of ρ and the inequality $(a+b)^\rho > a^\rho + b^\rho$ for $a, b > 0$ and $\rho > 1$. This is a contradiction, hence we have $\mathbb{E}|X(\alpha)|^\rho = \infty$. ∎

Zolotarev distance between $X_{n,k}/\mu_{n,k}$ and $X(\alpha)$.

Theorem 4. *If $0 < \alpha < 2$, then*

$$\zeta_2\left(\frac{X_{n,k}}{\mu_{n,k}}, X(\alpha)\right) = O\left(\frac{R(n)+1}{\lambda_n}\right).$$

If $2 \le \alpha < e$, then

$$\zeta_s\left(\frac{X_{n,k}}{\mu_{n,k}}, X(\alpha)\right) \to 0,$$

where s is specified as above.

In particular, this theorem implies the convergence in distribution of $X_{n,k}/\mu_{n,k}$ for $0 < \alpha < e$ and proves the first part of Theorem 1.

Convergence rate of the factors in (20).

Lemma 4. *With s and $R(n)$ specified as above, we have*

$$\left\| \frac{\mu_{I_n,k-1}}{\mu_{n,k}} - \alpha U^\alpha \right\|_s + \left\| \frac{\mu_{n-I_n,k}}{\mu_{n,k}} - (1-U)^\alpha \right\|_s = O\left(\frac{R(n)+1}{\lambda_n} \right).$$

Proof. We consider only the L_s-norm of $\mu_{I_n,k-1}/\mu_{n,k} - \alpha U^\alpha$, the other part being similar. By (3), we have

$$\mu_{n,k} = \frac{\mathsf{s}(n,k+1)}{(n-1)!} = \frac{\lambda_n^k}{k!} H(n,k),$$

where

$$H(n,k) = \frac{1}{\Gamma(1+\alpha_{n,k})} + O\left(\frac{1}{\lambda_n}\right), \tag{22}$$

the O-term holding uniformly for $1 \le k \le K\lambda_n$. Then we decompose the ratio $\mu_{I_n,k-1}/\mu_{n,k}$ into three parts

$$\frac{\mu_{I_n,k-1}}{\mu_{n,k}} = \frac{k}{\lambda_n}\left(\frac{\log I_n}{\lambda_n}\right)^{k-1} \frac{H(I_n,k-1)}{H(n,k)} =: F_n^{[1]} F_n^{[2]} F_n^{[3]}. \tag{23}$$

We first show that

$$|F_n^{[1]} - \alpha| + \| F_n^{[2]} - U^\alpha \|_{4s} + \| F_n^{[3]} - 1 \|_{4s} = O\left(\frac{R(n)+1}{\lambda_n}\right).$$

These estimates imply that $\|F_n^{[2]}\|_{4s}, \|F_n^{[3]}\|_{4s} = O(1)$. Then, Hölder's inequality gives

$$\left\| \frac{\mu_{I_n,k-1}}{\mu_{n,k}} - \alpha U^\alpha \right\|_s = O\left(\frac{R(n)+1}{\lambda_n}\right).$$

First, we introduce the set $\mathcal{A} := \{I_n \le n^{\alpha/6}\}$. Note that $\mu_{n,k} = O(1)$ for $k \ge 3\lambda_n$. On the set $\mathcal{A}$, we have $k-1 = \alpha\lambda_n + R(n) - 1 \ge (\alpha/2)\lambda_n \ge (\alpha/2)\log I_n^{6/\alpha} = 3\log I_n$, for sufficiently large n; thus $\mu_{I_n,k-1} = O(1)$. On the other hand, since $\alpha < e$, the mean satisfies $\mu_{n,k} = \Omega(1)$; thus

$$\int_{\mathcal{A}} \left| \frac{\mu_{I_n,k-1}}{\mu_{n,k}} - \alpha U^\alpha \right|^{4s} d\mathbb{P} = O(\mathbb{P}(\mathcal{A})) = O(\mathbb{P}(I_n \le \sqrt{n})) = O(1/\sqrt{n}) = O(\lambda_n^{-4s}).$$

Thus we need only to consider the complement set $\mathcal{A}^c$.

Obviously, $F_n^{[1]} = k/\lambda_n = \alpha + O(R(n)/\lambda_n)$.

For $F_n^{[2]}$, we observe that for $x \le 0$ the expansion $(1+x/m)^m = e^x + O(e^{\vartheta x}/m)$ holds uniformly with $\vartheta < 1$. Thus, we obtain

$$\begin{aligned} F_n^{[2]} = \left(\frac{\log I_n}{\lambda_n}\right)^{k-1} &= \left(\frac{I_n}{n} + O\left(\frac{(I_n/n)^\vartheta}{\lambda_n}\right)\right)^{\alpha + (R(n)-1)/\lambda_n} \\ &= U^\alpha + O\left(\frac{R(n)(U^\alpha + U^{\alpha+\vartheta-1})\log U + U^{\alpha+\vartheta-1}}{\lambda_n}\right). \end{aligned}$$

Here, we may choose ϑ with $1-\alpha < \vartheta < 1$. Then $(U^{\alpha} + U^{\alpha+\vartheta-1}) \log U$ and $U^{\alpha+\vartheta-1}$ are both L_{4s}-integrable and the O-term in the last display is bounded above by $O((R(n)+1)/\lambda_n)$ in L_{4s}.

For the third factor in (23), we have

$$H(n,k) = \frac{1}{\Gamma(1+\alpha+R(n)/\lambda_n)} + O\left(\frac{1}{\lambda_n}\right) = \frac{1}{\Gamma(1+\alpha)} + O\left(\frac{R(n)+1}{\lambda_n}\right).$$

For $H(I_n, k-1)$, we restrict to the set $\mathcal{A}^c$. On $\mathcal{A}^c$, for n sufficiently large, we have $k-1 \le 12 \log I_n$, so the error in the expansion of $H(I_n, k-1)$ implied by (22) is uniformly $O(1/\log I_n) = O(1/\lambda_n)$. Thus we have

$$\begin{aligned} H(I_n, k-1) &= \frac{1}{\Gamma\left(1+\alpha+\frac{\alpha \log(n/I_n)+R(n)-1}{\log I_n}\right)} + O\left(\frac{1}{\log I_n}\right) \\ &= \frac{1}{\Gamma(1+\alpha)} + O\left(\frac{\log(n/I_n)+R(n)}{\lambda_n}\right). \end{aligned}$$

Since $\|\log(n/I_n)\|_{4s} \to \|\log U\|_{4s} < \infty$, the last error term is of order $O((R(n)+1)/\lambda_n)$ in L_{4s}. Collecting all estimates, we obtain $\|F_n^{[3]} - 1\|_{4s} = O((R(n)+1)/\lambda_n)$. ∎

Asymptotic transfer of the double-indexed recurrence (14). Consider the recurrence (14) with suitable initial conditions.

Lemma 5. *If*

$$b_{n,k} = O\left(((v\lambda_n)^{-1/2} n^v v^{-k})^w \cdot \frac{R(n)+1}{\lambda_n}\right) \qquad (1 < w \le 2),$$

uniformly for $1 \le k < n$, *where* $0 < v < v_0$, *then*

$$a_{n,k} = O\left(\frac{1}{w - v^{w-1}} ((v\lambda_n)^{-1/2} n^v v^{-k})^w \cdot \frac{R(n)+1}{\lambda_n}\right), \tag{24}$$

uniformly for $1 \le k < n$, *where* $0 < v < \min\{w^{1/(w-1)}, v_0\}$.

Proof. The proof is similar to that for Lemma 3 but slightly more complicated. By the exact expression for $a_{n,k}$ and the estimate for $b_{n,k}$, we have, for $0 < v < v_0$,

$$a_{n,k} - b_{n,k} = O\left(v^{-wk-w/2} \sum_{1\le j<n} \sum_{0\le r\le k} |k-r-\alpha\lambda_j| \lambda_j^{-1-w/2} j^{wv-1} v^{wr} [u^r](1+u) \prod_{j<\ell<n} \left(1+\frac{u}{\ell}\right)\right).$$

First, if $|k-\alpha\lambda_n| \ge \varepsilon\lambda_n$, then $|k-r-\alpha\lambda_j| = O(k+\lambda_n)$, so that (24) holds by the proof of Lemma 3. We assume now that $|k-\alpha\lambda_n| \le \varepsilon\lambda_n$. Split the sum in j into three parts

$$\begin{aligned} a_{n,k} - b_{n,k} = O\Bigg(v^{-wk-w/2} &\left(\sum_{1\le j<\delta n} + \sum_{\delta n \le j \le (1-\delta)n} + \sum_{(1-\delta)n < j < n} \right) \\ &\times \sum_{0\le r\le k} |k-r-\alpha\lambda_j| \lambda_j^{-1-w/2} j^{wv-1} v^{wr} [u^r](1+u) \prod_{j<\ell<n} \left(1+\frac{u}{\ell}\right) \Bigg), \end{aligned}$$

where $\delta \in (0,1)$ will be specified later. An analysis similar to the proof of Lemma 3 gives

$$a_{n,k} - b_{n,k} = O\left(\frac{(v\lambda_n)^{-w/2}}{w - v^{w-1}} v^{-wk} n^{wv} \left(\delta^{wv-v^w} + \frac{|k-\alpha\lambda_n|+1}{\lambda_n} + \delta\right)\right),$$

where $0 < v < \min\{w^{1/(w-1)}, v_0\}$. Taking $\delta := ((R(n)+1)/\lambda_n)^{1/(wv-v^w)}$ yields (24). ∎

An inequality between ζ_s- and ℓ_s-distances.

Lemma 6. *For $1 < w \le 2$ and $M > 0$, there is a constant $K > 0$ such that*

$$\zeta_w(X, Y) \le K(\ell_w(X, Y) \vee \ell_w^{w-1}(X, Y)), \tag{25}$$

for all pairs $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_w(1)$ with $\|X\|_w, \|Y\|_w \le M$.

Proof. We start from the inequality (see Theorem 3, Zolotarev, 1976)

$$\zeta_w(X, Y) \le \frac{1}{w}\left(2\beta_w(X, Y) + 2^{w-1}\beta_w^{w-1}(X, Y)(\|X\|_w^w \wedge \|Y\|_w^w)^{2-w}\right),$$

for $1 < w \le 2$, where β_w denotes the difference pseudo-moment

$$\beta_w(\nu_1, \nu_2) := \inf\left\{\mathbb{E}\left||X|^{w-1}X - |Y|^{w-1}Y\right| : \mathcal{L}(X) = \nu_1, \mathcal{L}(Y) = \nu_2\right\} \qquad (w > 1),$$

with $\nu_1, \nu_2 \in \mathcal{M}_w$. From $\left||x|^{w-1}x - |y|^{w-1}y\right| \le w(|x|^{w-1} \vee |y|^{w-1})|x - y|$ and Hölder's inequality, it follows that

$$\beta_w(X, Y) \le w\,(\mathbb{E}|X|^w + \mathbb{E}|Y|^w)^{(w-1)/w}\,\ell_w(X, Y),$$

which implies the desired inequality. ∎

Proof of Theorem 4. We introduce a hybrid quantity

$$\Xi_n := \frac{\mu_{I_n,k-1}}{\mu_{n,k}}X(\alpha) + \frac{\mu_{n-I_n,k}}{\mu_{n,k}}X^*(\alpha),$$

where $X(\alpha), X^*(\alpha), I_n$ are independent and $X(\alpha), X^*(\alpha)$ identically distributed. Since $\mathcal{L}(X(\alpha)), \mathcal{L}(\bar{X}_{n,k})$, $\mathcal{L}(\Xi_n) \in \mathcal{M}_s(1)$, the ζ_s-distances between these quantities are finite. For simplicity, write $h_{n,k} := \zeta_s(\bar{X}_{n,k}, X(\alpha))$. By triangle inequality

$$h_{n,k} \le \zeta_s(\bar{X}_{n,k}, \Xi_n) + \zeta_s(\Xi_n, X(\alpha)).$$

Note that ζ_s is ideal of order s. Thus

$$\begin{aligned}
\zeta_s(\bar{X}_{n,k}, \Xi_n) &= \zeta_s\left(\frac{\mu_{I_n,k-1}}{\mu_{n,k}}\bar{X}_{I_n,k-1} + \frac{\mu_{n-I_n,k}}{\mu_{n,k}}\bar{X}^*_{n-I_n,k}, \frac{\mu_{I_n,k-1}}{\mu_{n,k}}X(\alpha) + \frac{\mu_{n-I_n,k}}{\mu_{n,k}}X^*(\alpha)\right) \\
&\le \frac{1}{n-1}\sum_{1\le j<n}\zeta_s\left(\frac{\mu_{j,k-1}}{\mu_{n,k}}\bar{X}_{j,k-1} + \frac{\mu_{n-j,k}}{\mu_{n,k}}\bar{X}^*_{n-j,k}, \frac{\mu_{j,k-1}}{\mu_{n,k}}X(\alpha) + \frac{\mu_{n-j,k}}{\mu_{n,k}}X^*(\alpha)\right) \\
&\le \frac{1}{n-1}\sum_{1\le j<n}\left(\left(\frac{\mu_{j,k-1}}{\mu_{n,k}}\right)^s h_{j,k-1} + \left(\frac{\mu_{n-j,k}}{\mu_{n,k}}\right)^s h_{n-j,k}\right).
\end{aligned}$$

We now show that

$$\zeta_s(\Xi_n, X(\alpha)) = O\left(D(n)^{s-1}\right), \tag{26}$$

where $D(n) := (R(n) + 1)/\lambda_n$.

First, by Lemma 4,

$$\|\Xi_n\|_s \leq \left(\left\| \frac{\mu_{I_n,k-1}}{\mu_{n,k}} \right\|_s + \left\| \frac{\mu_{n-I_n,k}}{\mu_{n,k}} \right\|_s \right) \|X(\alpha)\|_s$$
$$\to (\alpha \|U^\alpha\|_s + \|(1-U)^\alpha\|_s) \|X(\alpha)\|_s,$$

which implies that $\|\Xi_n\|_s$ is uniformly bounded for all n. Since $\mathcal{L}(X(\alpha)) \in \mathcal{M}_s(1)$, there is an $M > 0$ such that $\|X(\alpha)\|_s, \|\Xi_n\|_s \leq M$ for all n. We apply Lemma 6 to bound the ζ_s-distance, which gives

$$\zeta_s(\Xi_n, X(\alpha)) \leq K(\ell_s(\Xi_n, X(\alpha)) \vee \ell_s^{s-1}(\Xi_n, X(\alpha))).$$

By Lemma 4

$$\ell_s(\Xi_n, X(\alpha)) \leq \left(\left\| \frac{\mu_{I_n,k-1}}{\mu_{n,k}} - \alpha U^\alpha \right\|_s + \left\| \frac{\mu_{n-I_n,k}}{\mu_{n,k}} - (1-U)^\alpha \right\|_s \right) \|X(\alpha)\|_s = O\left(D(n)\right).$$

This proves (26).

Collecting the estimates, we obtain

$$h_{n,k} \leq \frac{1}{n-1} \sum_{1 \leq j < n} \left(\left(\frac{\mu_{j,k-1}}{\mu_{n,k}} \right)^s h_{j,k-1} + \left(\frac{\mu_{n-j,k}}{\mu_{n,k}} \right)^s h_{n-j,k} \right) + O\left(D(n)^{s-1}\right).$$

Thus, $h_{n,k} = O(a_{n,k}\mu_{n,k}^{-s})$, where $a_{n,k}$ satisfies (14) with

$$b_{n,k} = O\left(\mu_{n,k}^s D(n)^{s-1}\right),$$

and suitable initial conditions. Theorem 4 then follows from applying the different types of asymptotic transfer given in Lemmas 3 and 5. ∎

Remark. Note that the proof of Theorem 4 also yields a rate of convergence of order $O(((R(n)+1)/\lambda_n)^{s-1})$ for ζ_s for the range $2 \leq \alpha < e$.

Recently, S. Janson (private communication) showed that Lemma 6 also holds with (25) there replaced by

$$\zeta_w(X, Y) \leq K\ell_w(X, Y).$$

This inequality leads to an improvement of the error term in Theorem 4 for the range $2 \leq \alpha < e$ to $O((R(n)+1)/\lambda_n)$.

4 Asymptotics of moments

We prove in this section the moment estimate (6) whose proof is more involved than the asymptotic transfer in Lemma 3. The idea is to first derive a crude bound for higher moments of $X_{n,k}$, which holds uniformly for $1 \leq k < n$. Then a more refined analysis leads to (6).

Note that the m-th factorial moments of $X_{n,k}$ and the m-th moments are asymptotically equivalent when $\mu_{n,k} \to \infty$, or roughly when $\alpha < e$.

A uniform estimate for higher moments. For convenience, define $\varphi_1(v) = 1$ and

$$\varphi_m(v) := \frac{1}{m - v^{m-1}} \qquad (m \ge 2).$$

We now prove by induction that

$$A_{n,k}^{(m)} = O\left(\varphi_m(v)\left((v\lambda_n)^{-1/2} v^{-k} n^v\right)^m\right) \qquad (m \ge 1), \tag{27}$$

uniformly for $1 \le k < n$, where $0 < v < m^{1/(m-1)}$.

Obviously, (27) holds for $m = 1$ by (17). By (13) and induction, we have for $0 < v < (m-1)^{1/(m-2)}$

$$\begin{aligned}
B_{n,k}^{(m)} &= O\Bigg(\sum_{1 \le h < m} \binom{m}{h} \varphi_h(v)\varphi_{m-h}(v) \\
&\qquad \times n^{-1} \sum_{1 \le j < n} \left((v\lambda_j)^{-1/2} v^{-k+1} j^v\right)^h \left((v\lambda_{n-j})^{-1/2} v^{-k} (n-j)^v\right)^{m-h}\Bigg) \\
&= O\left(\varphi_{m-1}(v) v^{-km} n^{-1} \sum_{\substack{1 \le h < m \\ 1 \le j < n}} j^{hv} (n-j)^{(m-h)v} (v\lambda_j)^{-h/2} (v\lambda_{n-j})^{-(m-h)/2}\right) \\
&= O\left(\varphi_{m-1}(v)(v\lambda_n)^{-m/2} v^{-km} n^{mv}\right),
\end{aligned} \tag{28}$$

uniformly for $1 \le k < n$.

By (15),

$$A_{n,k}^{(m)} = B_{n,k}^{(m)} + \sum_{1 \le j < n} \sum_{0 \le r \le k} \frac{B_{j,k-r}^{(m)}}{j} [u^r](u+1) \prod_{j < \ell < n} \left(1 + \frac{u}{\ell}\right). \tag{29}$$

Substituting the estimate (28) into (29) gives for $0 < v < m^{1/(m-1)}$

$$\begin{aligned}
A_{n,k}^{(m)} &= O\left(B_{n,k}^{(m)} + v^{-km} \sum_{1 \le j < n} (v\lambda_j)^{-m/2} j^{mv-1} \sum_{0 \le r \le k} v^{rm} [u^r](1+u) \prod_{j < \ell < n} \left(1 + \frac{u}{\ell}\right)\right) \\
&= O\left(B_{n,k}^{(m)} + \varphi_m(v)(v\lambda_n)^{-m/2} n^{mv} v^{-km}\right),
\end{aligned}$$

similar to the proof of Lemma 3. This proves (27).

Note that when $\alpha \le m^{1/(m-1)} - \varepsilon$, the optimal choice of v in (27) minimizing $n^v v^{-k}$ is $v = \alpha_{n,k}$, which yields the estimate $A_{n,k}^{(m)} = O(\lambda_n^k / k!)$, uniformly in k. When $\alpha \ge m^{1/(m-1)} - \varepsilon$, the optimal choice is then $v = m^{1/(m-1)} - \varepsilon$. This says that the asymptotic behavior of $A_{n,k}^{(m)}$ when $\alpha < m^{1/(m-1)}$ is very different from that when $\alpha \ge m^{1/(m-1)}$. More precise estimates can be derived, but they are not needed here; see Drmota and Hwang (2005a) for asymptotic approximations to the variance (covering all ranges).

Asymptotics of $A_{n,k}^{(m)}$. Since the case $\alpha = 0$ will be treated separately, we assume throughout this section that $\alpha > 0$. We refine the above inductive argument and show that

$$A_{n,k}^{(m)} \sim \nu_m(\alpha)\mu_{n,k}^m \sim \nu_m(\alpha)\left(\frac{\lambda_n^k}{\Gamma(1+\alpha)k!}\right)^m, \tag{30}$$

for each $m \geq 1$ and $k/\lambda_n \to \alpha < m^{1/(m-1)}$, where $\nu_m(\alpha)$ denotes the moment sequence of $X(\alpha)$ given in (7). This will prove the moment convergence part of Theorem 1.

Note that by (3), (30) holds for $m = 1$ with $\nu_1(\alpha) = 1$. Assume that (30) holds for all $A_{n,k}^{(i)}$ with $i < m$. We split the right-hand side of (29) into three parts

$$\begin{aligned}
A_{n,k}^{(m)} &= B_{n,k}^{(m)} + \sum_{0\le r\le k}\left(\sum_{1\le j<\varepsilon n} + \sum_{\varepsilon n\le j\le(1-\varepsilon)n} + \sum_{(1-\varepsilon)n<j<n}\right)\frac{B_{j,k-r}^{(m)}}{j}[u^r](u+1)\prod_{j<\ell<n}\left(1+\frac{u}{\ell}\right)\\
&=: B_{n,k}^{(m)} + A_{n,k}^{(m)}[1] + A_{n,k}^{(m)}[2] + A_{n,k}^{(m)}[3].
\end{aligned}$$

By the same proof used for Lemma 3, we have

$$\begin{aligned}
A_{n,k}^{(m)}[1] &= O\left(\varepsilon^{mv-v^m}\varphi_m(v)\lambda_n^{-(m+1)/2}n^{mv}v^{-km}\right),\\
A_{n,k}^{(m)}[3] &= O\left(\varepsilon\varphi_m(v)\lambda_n^{-(m+1)/2}n^{mv}v^{-km}\right).
\end{aligned}$$

Letting $\varepsilon \to 0$, we see that, by (27),

$$A_{n,k}^{(m)}[1] + A_{n,k}^{(m)}[3] = o(A_{n,k}^{(m)}).$$

Asymptotics of $A_{n,k}^{(m)}$: the dominant terms. We start by showing that for $0 < \alpha < (m-1)^{1/(m-2)}$

$$B_{n,k}^{(m)} \sim \nu_m^*(\alpha)\left(\frac{\lambda_n^k}{\Gamma(1+\alpha)k!}\right)^m \qquad (m \geq 2), \tag{31}$$

where

$$\nu_m^*(\alpha) := \sum_{1\le h<m}\binom{m}{h}\nu_h(\alpha)\nu_{m-h}(\alpha)\alpha^h\int_0^1 u^{h\alpha}(1-u)^{(m-h)\alpha}\,\mathrm{d}u.$$

By (13), induction and (30), we have, for $0 < \alpha < (m-1)^{1/(m-2)}$,

$$\begin{aligned}
B_{n,k}^{(m)} &\sim \sum_{1\le h<m}\binom{m}{h}\nu_h(\alpha)\nu_{m-h}(\alpha)\frac{1}{n}\sum_{\varepsilon n\le j\le(1-\varepsilon)n}\left(\frac{\lambda_j^{k-1}}{\Gamma(1+\alpha)(k-1)!}\right)^h\left(\frac{\lambda_{n-j}^k}{\Gamma(1+\alpha)k!}\right)^{m-h}\\
&\sim \left(\frac{\lambda_n^k}{\Gamma(1+\alpha)k!}\right)^m\sum_{1\le h<m}\binom{m}{h}\nu_h(\alpha)\nu_{m-h}(\alpha)\frac{1}{n}\sum_{\varepsilon n\le j\le(1-\varepsilon)n}\alpha^h\left(\frac{j}{n}\right)^{kh/\lambda_n}\left(1-\frac{j}{n}\right)^{k(m-h)/\lambda_n},
\end{aligned}$$

which proves (31). The errors introduced for terms with $j < \varepsilon n$ and for $j \geq (1-\varepsilon)n$ can be easily bounded by using (27).

To evaluate $A_{n,k}^{(m)}[2]$, we first observe that

$$\begin{aligned}\prod_{j<\ell<n}\left(1+\frac{u}{\ell}\right) &= \exp\left(u\sum_{j<\ell<n}\ell^{-1}+O\left(\frac{|u|^2}{j}\right)\right)\\ &= (n/j)^u\left(1+O\left(|u|^2 j^{-1}\right)\right),\end{aligned}$$

uniformly for finite complex u and $j\to\infty$. It follows that

$$[u^r]\prod_{j<\ell<n}\left(1+\frac{u}{\ell}\right)=\frac{(\log(n/j))^r}{r!}\left(1+O\left(\frac{r^2}{j}\right)\right),$$

uniformly for $\varepsilon n\le j\le(1-\varepsilon)n$ and $0\le r\le k=o(\sqrt{j})$. Consequently, by (28) and (31),

$$\begin{aligned}A_{n,k}^{(m)}[2] &\sim \nu_m^*(\alpha)\left(\frac{\lambda_n^k}{\Gamma(1+\alpha)k!}\right)^m\sum_{\varepsilon n\le j\le(1-\varepsilon)n} j^{-1}(j/n)^{m\alpha}\\ &\qquad\times\sum_{r\ge0}\alpha^{mr}\left(\frac{(\log(n/j))^{r-1}}{(r-1)!}+\frac{(\log(n/j))^r}{r!}\right)\\ &\sim \nu_m^*(\alpha)(\alpha^m+1)\left(\frac{\lambda_n^k}{\Gamma(1+\alpha)k!}\right)^m\int_\varepsilon^{1-\varepsilon}x^{m\alpha-\alpha^m-1}\,\mathrm{d}x.\end{aligned}$$

Letting $\varepsilon\to0$, we then obtain, by (29), that

$$\begin{aligned}A_{n,k}^{(m)} &\sim \nu_m^*(\alpha)\left(1+(\alpha^m+1)\int_0^1 x^{m\alpha-\alpha^m-1}\,\mathrm{d}x\right)\left(\frac{\lambda_n^k}{\Gamma(1+\alpha)k!}\right)^m\\ &= \nu_m^*(\alpha)\frac{m\alpha+1}{m\alpha-\alpha^m}\left(\frac{\lambda_n^k}{\Gamma(1+\alpha)k!}\right)^m,\end{aligned}$$

where

$$\begin{aligned}\nu_m^*(\alpha)\frac{m\alpha+1}{m\alpha-\alpha^m} &= \frac{1}{m-\alpha^{m-1}}\sum_{1\le h<m}\binom{m}{h}\nu_h(\alpha)\nu_{m-h}(\alpha)\alpha^{h-1}\frac{\Gamma(h\alpha+1)\Gamma((m-h)\alpha+1)}{\Gamma(m\alpha+1)}\\ &= \nu_m(\alpha),\end{aligned}$$

for $m\ge2$, by (7). This completes the proof of (29) and thus Theorem 1 (*ii*). ∎

Moment convergence (6). Convergence of all moments implies convergence in distribution if the moment sequence (7) uniquely characterizes the distribution. By considering $\bar\nu_m(\alpha):=\nu_m(\alpha)\Gamma(m\alpha+1)/m!$, we easily obtain by induction that $\bar\nu_m(\alpha)=O(K^m)$ for $\alpha\in[0,1]$ (see Hwang and Neininger, 2002), and thus convergence in distribution of $X_{n,k}/\mu_{n,k}$ follows from (6) when $\alpha\in[0,1]$.

5 The central range $\alpha=1$

We prove Theorem 2 in this section. The proof proceeds essentially along the same line as we did above but with one major difference: we consider central moments instead of factorial moments. This minor step is crucial in dealing with the cancellations involved in the asymptotics of higher central moments. For simplicity, the case when $|t_{n,k}| \to \infty$ and $t_{n,k} = o(\lambda_n)$ is first analyzed; then the same method of proof is extended to the case when $t_{n,k} = O(1)$. Justifications of the error terms are similar to those for $A_{n,k}^{(m)}$ given above but become more complicated.

Recurrence of central moments. Consider $\bar{P}_{n,k}(y) := \mathbb{E}(e^{(X_{n,k}-\mu_{n,k})y}) = P_{n,k}(e^y)e^{-\mu_{n,k}y}$; see (12). Then we have the recurrence

$$\bar{P}_{n,k}(y) = \frac{1}{n-1}\sum_{1\le j<n} \bar{P}_{j,k-1}(y)\bar{P}_{n-j,k}(y)e^{\Delta_{n,k}(j)y} \qquad (n \ge 2; k \ge 1),$$

where

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

and $\bar{P}_{n,0}(y) = \bar{P}_{1,k}(y) = 1$ for $n,k \ge 1$.

Let now $P_{n,k}^{(m)} := \bar{P}_{n,k}^{(m)}(0)$ denote the m-th central moment of $X_{n,k}$. Then $P_{n,k}^{(1)} \equiv 0$ and for $m \ge 2$

$$P_{n,k}^{(m)} = \frac{1}{n-1}\sum_{1\le j<n}\left(P_{j,k-1}^{(m)} + P_{j,k}^{(m)}\right) + Q_{n,k}^{(m)} \qquad (n \ge 2; k \ge 1), \tag{32}$$

where

$$Q_{n,k}^{(m)} := \sum_{\substack{a+b+c=m\\0\le a,b<m\\0\le c\le m}} \binom{m}{a,b,c}\frac{1}{n-1}\sum_{1\le j<n} P_{j,k-1}^{(a)}P_{n-j,k}^{(b)}\Delta_{n,k}^{c}(j)$$

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

Outline of the proof of Theorem 2. Similar to the proof of (30), we divide the proof of Theorem 2 into three main steps.

- We first derive a uniform estimate for $\Delta_{n,k}(j)$ for $1 \le j,k < n$, which then implies a uniform bound for $P_{n,k}^{(m)}$ for $1 \le k < n$. This bound is sufficient for our uses except when $|k-\lambda_n| = o(\sqrt{\lambda_n})$.
- We then derive a second estimate for $\Delta_{n,k}(j)$ uniformly valid for $k \sim \lambda_n$. This in turn implies a tight bound for $P_{n,k}^{(m)}$ when $k \sim \lambda_n$, and an asymptotic approximation to $P_{n,k}^{(m)}$ when $1 \ll |t_{n,k}| = o(\lambda_n)$.
- A finer estimate for $\Delta_{n,k}(j)$ is needed to deal with the case when $t_{n,k} = O(1)$.

An integral representation for $\Delta_{n,k}(j)$. By (2),

$$\mu_{n,k} = [u^k]\frac{n^u}{\Gamma(u+1)}\left(1 + O\left(n^{-1}\right)\right).$$

Then

$$\Delta_{n,k}(j) = \frac{1}{2\pi i}\oint_{|u|=v} u^{-k-1}n^u\phi(u,j/n)\left(1 + O(j^{-1} + (n-j)^{-1})\right)\,\mathrm{d}u, \tag{33}$$

uniformly for $1 \le j < n$ (when j or $n-j$ is bounded, the O-term becoming $O(1)$ instead of $o(1)$), where

$$\phi(u,x) := \frac{(1-x)^u + ux^u - 1}{\Gamma(u+1)}.$$

Here and throughout this section, we take $v = 1 + o(1)$ since $k \sim \lambda_n$.

A uniform estimate for $\Delta_{n,k}(j)$. Since $\phi(1,x)=0$, we have

$$|\phi(u,x)| = O(|u-1|) \qquad (x \in [0,1]).$$

Substituting this estimate into (33) gives

$$\begin{aligned}\Delta_{n,k}(j) &= O\left(v^{-k}n^{v}\int_{-\pi}^{\pi}\left|ve^{i\theta}-1\right|n^{-v(1-\cos\theta)}\mathrm{d}\theta\right)\\ &= O\left((|v-1|+\lambda_n^{-1/2})\lambda_n^{-1/2}v^{-k}n^{v}\right),\end{aligned} \tag{34}$$

uniformly for $1 \le j,k < n$.

A uniform estimate for $P_{n,k}^{(m)}$. From the recurrence (32) and the estimate (34), we deduce, by an induction similar to that used for (27), that

$$Q_{n,k}^{(m)}, P_{n,k}^{(m)} = O\left((|v-1|^m + \lambda_n^{-m/2})\left(\lambda_n^{-1/2}v^{-k}n^{v}\right)^m\right) \qquad (m \ge 2), \tag{35}$$

uniformly for $1 \le k < n$. This bound is however not tight when $|k-\lambda_n| = o(\sqrt{\lambda_n})$, the reason being simply that v is not properly chosen to minimize the error term (the first $\lambda_n^{-1/2}$) in (34).

A finer estimate than (34). For a more precise estimate than (34), we use the two-term Taylor expansion

$$\phi(u,x) = \phi_u'(1,x)(u-1) + O(|u-1|^2),$$

where $\phi_u'(1,x) = x + x\log x + (1-x)\log(1-x)$, which leads to

$$\begin{aligned}\Delta_{n,k}(j) &= \phi_u'\left(1,\frac{j}{n}\right)(k-\lambda_n)\frac{\lambda_n^{k-1}}{k!}\left(1+O(j^{-1}+(n-j)^{-1})\right)\\ &\quad + O\left((|v-1|^2+\lambda_n^{-1})\lambda_n^{-1/2}v^{-k}n^{v}\right).\end{aligned} \tag{36}$$

Taking $v = k/\lambda_n$ gives

$$\Delta_{n,k}(j) = O\left((|k-\lambda_n|+1)\frac{\lambda_n^{k-1}}{k!}\right). \tag{37}$$

This bound holds uniformly for $k \sim \lambda_n$ and $1 \le j < n$ since $\phi_u'(1,x) = O(x|\log x|)$ as $x \to 0^+$.

A uniform bound for $P_{n,k}^{(m)}$ when $k \sim \lambda_n$. From (37), we deduce, again by induction, that

$$Q_{n,k}^{(m)}, P_{n,k}^{(m)} = O\left((|k-\lambda_n|^m+1)\left(\frac{\lambda_n^{k-1}}{k!}\right)^m\right) \qquad (m \ge 2), \tag{38}$$

uniformly for $k \sim \lambda_n$. The proof differs slightly from that for (30) in that we split all sums of the form $\sum_{1\le j<n}$ into three parts

$$\sum_{1\le j<n} = \sum_{1\le j<n/\lambda_n^m} + \sum_{n/\lambda_n^m \le j \le n-n/\lambda_n^m} + \sum_{n-n/\lambda_n^m<j<n},$$

and then apply (38) and (37) to the middle sum, and (35) to the remaining two sums.

Asymptotics of $P_{n,k}^{(m)}$ when $|t_{n,k}| \to \infty$ and $t_{n,k} = o(\lambda_n)$. In this case, the estimate (36) has the form

$$\Delta_{n,k}(j) \sim \phi_u'\left(1, \frac{j}{n}\right) t_{n,k} \frac{\lambda_n^{k-1}}{k!}, \tag{39}$$

uniformly in k when $\varepsilon n \leq j \leq (1-\varepsilon)n$. Then we show that

$$P_{n,k}^{(m)} \sim g_m \left(t_{n,k} \frac{\lambda_n^{k-1}}{k!}\right)^m \qquad (m \geq 1), \tag{40}$$

where $g_0 = 1$, $g_1 = 0$ and for $m \geq 2$

$$g_m = \frac{m+1}{m-1} \sum_{\substack{a+b+c=m \\ 0\leq a,b<m \\ 0\leq c\leq m}} \binom{m}{a,b,c} g_a g_b \int_0^1 x^a (1-x)^b \phi_u'(1,x)^c \, \mathrm{d}x. \tag{41}$$

Equivalently, this can be written as

$$g_m = \sum_{\substack{a+b+c=m \\ 0\leq a,b,c\leq m}} \binom{m}{a,b,c} g_a g_b \int_0^1 x^a (1-x)^b \phi_u'(1,x)^c \, \mathrm{d}x.$$

In particular,

$$g_2 = 3 \int_0^1 \phi_u'(1,x)^2 \, \mathrm{d}x = 2 - \frac{\pi^2}{6}.$$

The inductive proof is almost the same as that for $A_{n,k}^{(m)}$, with the factor $(k-\lambda_n)^m$ handled by direct expansion and then estimated term by term. Also we need to split sums of the form $\sum_{1\leq j<n}$ into five parts

$$\sum_{1\leq j<n} = \sum_{1\leq j<n/\lambda_n^m} + \sum_{n/\lambda_n^m\leq j<\varepsilon n} + \sum_{\varepsilon n\leq j\leq (1-\varepsilon)n} + \sum_{(1-\varepsilon)n<j\leq n-n/\lambda_n^m} + \sum_{n-n/\lambda_n^m<j<n},$$

and then apply (40) to the middle sum, and the two estimates (35) and (38) to the other four sums.

The moment sequence (41) is easily checked to have the property of uniquely characterizing the distribution; see Hwang (2005) for similar details.

This proves the first part of Theorem 2.

The periodic case when $t_{n,k} = O(1)$. In this case, we need a more precise expansion than (39) as follows.

$$\Delta_{n,k}(j) \sim \frac{\lambda_n^{k-1}}{k!} \left(\phi_u'\left(1, \frac{j}{n}\right) t_{n,k} - \frac{1}{2}\phi_{uu}''\left(1, \frac{j}{n}\right)\right), \tag{42}$$

uniformly for $j/n \in [\varepsilon, 1-\varepsilon]$ and $k \sim \lambda_n$, where

$$\phi_{uu}''(1,x) = (x \log x + (1-x)\log(1-x))^2 - 2(1-\gamma)\phi_u'(1,x).$$

This is proved by expanding more terms of $\phi(u,x)$ at $u = 1$ and then estimating the error terms (see Hwang, 1995 for similar details).

With the approximation (42), we first prove that for $m \geq 0$

$$\mathbb{E}(X_{n,k} - \mu_{n,k})^m = P_{n,k}^{(m)} \sim p_m(t_{n,k}) \left(\frac{\lambda_n^{k-1}}{k!}\right)^m, \tag{43}$$

where $p_m(t_{n,k})$ is a polynomial in $t_{n,k}$ of degree m with $p_0(t_{n,k}) = 1$ and $p_1(t_{n,k}) = 0$. This will imply that for $k = \lfloor \lambda_n \rfloor + \ell$, where $\ell \in \mathbb{Z}$,

$$\mathbb{E}\left(\frac{X_{n,k} - \mu_{n,k}}{\lambda_n^{k-1}/k!}\right)^m \sim p_m(\ell - \{\lambda_n\}),$$

for $m \geq 0$, where $\{\lambda_n\}$ denotes the fractional part of λ_n. Then we apply an argument based on the Frechet-Shohat moment convergence theorem similar to that used in Chern and Hwang (2001a) to prove that $(X_{n,k} - \mu_{n,k})/(\lambda_n^{k-1}/k!)$ does not converge to a fixed limit law. The proof for $(X_{n,k} - \mu_{n,k})/\sqrt{\mathbb{V}(X_{n,k})}$ is similar.

To prove (43), we use again induction. Assume $m \geq 2$. Then a similar analysis as above leads to

$$Q_{n,k}^{(m)} \sim q_m(t_{n,k}) \left(\frac{\lambda_n^{k-1}}{k!}\right)^m,$$

where $q_m(t)$ is a polynomial of degree m defined by

$$\begin{aligned} q_m(t_{n,k}) := \sum_{\substack{a+b+c=m \\ 0 \leq a,b<m \\ 0 \leq c \leq m}} \binom{m}{a,b,c} \int_0^1 y^a(1-y)^b \\ \times p_a(t_{n,k} - 1 - \log y)\, p_b(t_{n,k} - \log(1-y)) \left(\phi_u'(1,y)t_{n,k} - \frac{1}{2}\phi_{uu}''(1,y)\right)^c \mathrm{d}y. \end{aligned}$$

Then by (32), we deduce that for $m \geq 2$

$$\begin{aligned} P_{n,k}^{(m)} \left(\frac{\lambda_n^{k-1}}{k!}\right)^{-m} \sim q_m(t_{n,k}) + \int_0^1 x^{m-1} \sum_{r \geq 0} \frac{\log(1/x)^r}{r!} \\ \times (q_m(t_{n,k} - r - 1 - \log x) + q_m(t_{n,k} - r - \log x))\, \mathrm{d}x, \end{aligned}$$

the infinite series on the right-hand side being convergent since q_m is a polynomial of degree m. This proves (43) and the second part of Theorem 2.

Note that by induction

$$p_m(t) = q_m(t) + \int_0^1 x^m \left(p_m(t - 1 - \log x) + p_m(t - \log x)\right) \mathrm{d}x \qquad (m \geq 2).$$

Straightforward calculation of the integrals gives the expression (10) for $p_2(t_{n,k})$.

Extrema of $|\mathbb{E}(X_{n,k} - \mu_{n,k})^m|$. To prove the maximum order of $\mathbb{E}(X_{n,k} - \mu_{n,k})^m$, we consider two cases. First, when $|k - \lambda_n| \leq \lambda^{2/3}$, we apply (38), so that

$$\begin{aligned} \max_{|k-\lambda_n| \leq \lambda_n^{2/3}} |P_{n,k}^{(m)}| &= O\left(\lambda^{-3m/2} n^m \cdot \max_{|t_{n,k}| \leq \lambda_n^{2/3}} \left(t_{n,k}^m + 1\right) e^{-m t_{n,k}^2/(2\lambda_n)}\right) \\ &= O\left(\lambda^{-m} n^m\right), \end{aligned}$$

the maximum being reached when $t_{n,k} \sim \pm\sqrt{\lambda_n}$.

On the other hand, when $|k-\lambda_n| \geq \lambda^{2/3}$, we apply the estimate (35) and bound the maximum by the sum

$$\max_{|k-\lambda_n|\geq\lambda_n^{2/3}} |P_{n,k}^{(m)}| = O\left(|v-1|^m \lambda_n^{-m/2} n^{mv} \left(\sum_{k\leq\lambda_n-\lambda^{2/3}} + \sum_{k\geq\lambda+\lambda^{2/3}}\right) v^{-mk}\right).$$

Taking $v = 1-\lambda_n^{-1/3}$ in the first sum and $v = 1+\lambda_n^{-1/3}$ in the second, we obtain

$$\max_{|k-\lambda_n|\geq\lambda_n^{2/3}} |P_{n,k}^{(m)}| = O\left(\lambda_n^{1/3-5m/6} n^m e^{-m\lambda_n^{1/3}/2}\right).$$

Thus

$$\max_{1\leq k<n} |\mathbb{E}(X_{n,k}-\mu_{n,k})^m| = O\left(\lambda_n^{-m} n^m\right).$$

The proof for the minimum order is similar. This proves Corollary 5.

6 Asymptotic normality when $\alpha = 0$

The approach we use in this section relies on manipulating the recurrences of two sequences of polynomials defined from the bivariate generating functions $P_k(z,y) := \sum_n \mathbb{E}(y^{X_{n,k}})z^n$. It can not only be applied to prove Theorem 3 but also gives an alternative proof of the moment convergence part of Theorem 1.

Main steps. Let

$$\sigma_{n,k} := \sqrt{\frac{\lambda_n^{2k-1}}{(k-1)!^2(2k-1)}},$$

$X_{n,k}^* := (X_{n,k}-\lambda_n^k/k!)/\sigma_{n,k}$, and $\Lambda := \lambda_n/k$. The proof of Theorem 3 uses the following estimates.

Proposition 2. *The characteristic functions of $X_{n,k}^*$ satisfy the two estimates: (i)*

$$\left|\mathbb{E}(e^{X_{n,k}^* i\theta}) - e^{-\theta^2/2}\right| = O\left(e^{-\theta^2/2}\frac{|\theta|+|\theta|^3}{\sqrt{\Lambda}} + n^{-\varepsilon}\right), \tag{44}$$

uniformly for $|\theta| \leq \varepsilon\Lambda^{1/6}$; and (ii)

$$\mathbb{E}(e^{X_{n,k}^* i\theta}) = O(e^{-\theta^2/4} + n^{-\varepsilon}), \tag{45}$$

uniformly for $\varepsilon\Lambda^{1/6} \leq |\theta| \leq \varepsilon\sqrt{\Lambda}$.

Theorem 3 then follows from applying the Berry-Esseen smoothing inequality (see Petrov, 1975).

These estimates are derived by singularity analysis (see Flajolet and Odlyzko, 1990), starting from Cauchy's integral representation

$$\mathbb{E}(e^{X_{n,k} i\theta/\sigma_{n,k}}) = \frac{1}{2\pi i}\oint_{|z|=\varepsilon} z^{-n-1} P_k(z, e^{i\theta/\sigma_{n,k}})\,\mathrm{d}z.$$

We then need estimates for the generating functions P_k, and for that purpose, we introduce two sequences of polynomials and derive approximations to P_k via those for the two polynomials.

Two sequences of polynomials. By (12), the generating function P_k satisfies

$$\begin{cases} P_0(z,y) = 1 + \dfrac{yz}{1-z}, \\ P_k(z,y) = 1 + z\exp\left(\displaystyle\int_0^z \frac{P_{k-1}(t,y)-1}{t}\,\mathrm{d}t\right) \quad (k \geq 1). \end{cases}$$

It is more convenient to work with

$$Q_k(z,s) := \frac{P_k(z,e^s)-1}{z}.$$

Then

$$\begin{cases} Q_0(z,s) = \dfrac{e^s}{1-z}, \\ Q_k(z,s) = \exp\left(\displaystyle\int_0^z Q_{k-1}(t,s)\,\mathrm{d}t\right), \qquad (k \geq 1). \end{cases} \tag{46}$$

Now, write $L(z) := -\log(1-z)$. We define two sequences of polynomials V and W as follows.

$$\begin{aligned} Q_k(z,s) &:= \exp\left(\sum_{m\geq 0} \frac{V_{k,m}(L(z))}{m!}\, s^m\right) \\ &:= \frac{1}{1-z}\sum_{m\geq 0} \frac{W_{k,m}(L(z))}{m!}\, s^m. \end{aligned}$$

Lemma 7. *The two sequences of polynomials satisfy the recurrences*

$$\begin{cases} V_{k,m}(x) = \displaystyle\int_0^x W_{k-1,m}(t)\,\mathrm{d}t \qquad (k \geq 2), \\ W_{k,m}(x) = \dfrac{1}{m}\displaystyle\sum_{1\leq j\leq m} \binom{m}{j} j\, V_{k,j}(x) W_{k,m-j}(x) \qquad (m \geq 1), \end{cases} \tag{47}$$

where $V_{1,m} = x$ for $m \geq 0$ and $W_{k,0}(x) = 1$ for $k \geq 1$.

Proof. The first relation follows from (46) and the second from taking derivative with respect to s and then collecting the coefficient of s^m on both sides. ∎

Mean value and variance. We first rederive the mean and variance by such a VW-polynomial approach.

By (47) with $m = 1$, we obtain

$$V_{k,1}(x) = W_{k,1}(x) = \frac{x^k}{k!} \qquad (k \geq 1). \tag{48}$$

Consequently, with $x = L(z)$,

$$\mu_{n,k} = [z^n]\frac{z}{1-z}\cdot\frac{L^k(z)}{k!} = \frac{\mathrm{s}(n,k+1)}{(n-1)!},$$

which rederives (2). The asymptotic behavior of $\mu_{n,k}$ when $k = o(\lambda_n)$ is derived as follows.

$$\begin{aligned}
\mu_{n,k} &= [u^k]\frac{n^u}{\Gamma(1+u)}\left(1+O(n^{-1})\right)\\
&= \frac{\lambda_n^k}{k!}\sum_{0\le j\le k}\frac{k!}{(k-j)!\lambda_n^j}\cdot[u^j]\frac{1}{\Gamma(1+u)} + O\left(\frac{\lambda_n^k}{nk!}\right)\\
&\sim \frac{\lambda_n^k}{k!}.
\end{aligned}$$

For $m = 2$, we have, again by (47),

$$\begin{aligned}
V_{k,2}(x) &= \int_0^x W_{k-1,2}(t)\,\mathrm{d}t = \frac{x^{2k-1}}{(k-1)!^2(2k-1)} + \int_0^x V_{k-1,2}(t)\,\mathrm{d}t\\
&= \sum_{0\le j<k}\binom{2j}{j}\frac{x^{k+j}}{(k+j)!};
\end{aligned}\tag{49}$$

and then

$$W_{k,2}(x) = V_{k,2}(x) + V_{k,1}^2(x) = \sum_{0\le j\le k}\binom{2j}{j}\frac{x^{k+j}}{(k+j)!}.$$

Hence,

$$\begin{aligned}
\mathbb{E}(X_{n,k}^2) &= [z^n]\frac{z}{1-z}\cdot\sum_{0\le j\le k}\binom{2j}{j}\frac{L^{k+j}(z)}{(k+j)!} = \sum_{0\le j\le k}\binom{2j}{j}\frac{\mathsf{s}(n,k+j+1)}{(n-1)!}\\
&= \sum_{0\le j\le k}\binom{2j}{j}[u^{k+j}]\frac{n^u}{\Gamma(1+u)}\left(1+O(n^{-1})\right);
\end{aligned}$$

cf. Meir and Moon (1978) and van der Hofstad et al. (2001). Now, observe that for $k = o(\lambda_n)$

$$\binom{2k}{k}[u^{2k}]\frac{n^u}{\Gamma(1+u)} - \left([u^k]\frac{n^u}{\Gamma(1+u)}\right)^2 = O\left(\frac{k^2\lambda_n^{2k-2}}{k!^2}\right).$$

It follows that

$$\mathbb{V}(X_{n,k}) \sim \frac{\lambda_n^{2k-1}}{(k-1)!^2(2k-1)} \qquad (k = o(\lambda_n)),$$

which proves the variance estimate in Theorem 3.

This line of computations can be extended to higher moments. For example, a similar reasoning for $m = 3$ yields

$$\begin{aligned}
V_{k,3}(x) &= \int_0^x V_{k-1,3}(t)\,\mathrm{d}t + \int_0^x \left(3V_{k-1,2}(t)V_{k-1,1}(t) + V_{k-1,1}^3(t)\right)\mathrm{d}t\\
&= 3\sum_{0\le\ell<k}\sum_{0\le j<\ell}\binom{2j}{j}\binom{j+2\ell}{\ell}\frac{x^{k+j+\ell}}{(k+j+\ell)!} + \sum_{0\le j<k}\binom{3j}{j,j,j}\frac{x^{k+2j}}{(k+2j)!};
\end{aligned}$$

and

$$W_{k,3}(x) = 3\sum_{0\le\ell\le k}\sum_{0\le j<\ell}\binom{2j}{j}\binom{j+2\ell}{\ell}\frac{x^{k+j+\ell}}{(k+j+\ell)!} + \sum_{0\le j\le k}\binom{3j}{j,j,j}\frac{x^{k+2j}}{(k+2j)!},$$

which was used to compute $\mathbb{E}(X_{n,k}-\mu_{n,k})^3$ in Figure 1. However, the resulting expressions soon become very involved. Thus we focus directly on asymptotics of these polynomials and not on exact expressions.

Asymptotics of the V and W polynomials. First, by (48), we have

$$V_{k,1}(x) = W_{k,1}(x) \sim \frac{x^k}{k!} \qquad (x \in \mathbb{C}),$$

for $k = o(|x|)$.

Next, by (49), we have the following estimates for $k = o(x)$

$$\begin{aligned} V_{k,2}(x) &= \frac{x^{2k-1}}{(k-1)!^2(2k-1)}\left(1 + \sum_{1\le j\le k}\frac{2k-1}{2k-j}\prod_{1\le \ell<j}\left(\frac{k-\ell}{x}\cdot\frac{k-\ell}{2k-j-\ell}\right)\right) \\ &\sim \frac{x^{2k-1}}{(k-1)!^2(2k-1)}, \end{aligned}$$

and

$$W_{k,2}(x) = V_{k,2}(x) + V_{k,1}^2(x) \sim \frac{x^{2k}}{k!^2}.$$

The general pattern is as follows.

Lemma 8. *If $k = o(|x|)$, where $x \in \mathbb{C}$ is large, then*

$$\begin{cases} V_{k,m}(x) \sim \dfrac{x^{m(k-1)+1}}{(k-1)!^m(m(k-1)+1)}, \\ W_{k,m}(x) \sim \dfrac{x^{mk}}{k!^m}. \end{cases} \tag{50}$$

Proof. We use induction on m. We already proved (50) for $m = 1, 2$. Assume $m \ge 3$. By (47) and induction

$$\begin{aligned} V_{k,m}(x) &= \int_0^x W_{k-1,m}(t)\,\mathrm{d}t \\ &\sim \frac{1}{m}\sum_{1\le j<m}\binom{m}{j} j \int_0^x \frac{t^{j(k-2)+1}}{(k-2)!^j(j(k-2)+1)}\cdot\frac{t^{(k-1)(m-j)}}{(k-1)!^{m-j}}\,\mathrm{d}t + \int_0^x V_{k-1,m}(t)\,\mathrm{d}t \\ &\sim \frac{x^{(k-1)m+1}}{(k-1)!^m((k-1)m+1)} + \int_0^x V_{k-1,m}(t)\,\mathrm{d}t. \end{aligned}$$

Hence, by iteration,

$$\begin{aligned} V_{k,m}(x) &\sim \sum_{0\le j<k}\frac{(mj)!}{j!^m}\cdot\frac{x^{k+j(m-1)}}{(k+j(m-1))!} \\ &\sim \frac{x^{(k-1)m+1}}{(k-1)!^m((k-1)m+1)}. \end{aligned}$$

Moreover, by applying (47) and induction again

$$\begin{aligned} W_{k,m}(x) &\sim \frac{1}{m}\sum_{1\le j\le m}\binom{m}{j} j \frac{x^{j(k-1)+1}}{(k-1)!^j(j(k-1)+1)}\cdot\frac{x^{k(m-j)}}{k!^{m-j}} \\ &\sim \frac{x^{mk}}{k!^m}. \end{aligned}$$

This proves (50). ∎

Proof of Proposition 2. By Cauchy's formula, we have

$$\mathbb{E}\left(e^{X_{n,k} i\theta/\sigma_{n,k}}\right) = \frac{1}{2\pi i}\oint_{|z|=\varepsilon} z^{-n} Q_k\left(z, i\theta/\sigma_{n,k}\right) \mathrm{d}z.$$

We then deform the integration circle onto the left contour shown in Figure 3, where $\delta_n = \lambda_n^2/n$. For the larger circle, we have

$$\frac{1}{2\pi i}\int_{|z|=1+\delta_n/n} z^{-n} Q_k(z, i\theta/\sigma_{n,k})\,\mathrm{d}z = O\left(e^{-\lambda_n^2} \sup_{|z|=1+\delta_n/n} |Q_k(z, i\theta/\sigma_{n,k})|\right).$$

Now by the estimate

$$\sigma_{n,k} = O\left(\Lambda^{-1/2}\frac{\lambda_n^k}{k!}\right),$$

and (50), we have

$$V_{k,m}(\log(n/\omega_n))\sigma_{n,k}^{-m} = O\left(\Lambda^{-(m-2)/2}\right) \qquad (m \geq 1),$$

for any complex sequence ω_n satisfying $1 \ll |\omega_n| = O(\lambda_n^K)$. It follows that the contribution from the large circle is bounded above by

$$\begin{aligned}\frac{1}{2\pi i}\int_{|z|=1+\delta_n/n} z^{-n} Q_k(z, i\theta/\sigma_{n,k})\,\mathrm{d}z &= O\left(n\lambda_n^{-2}e^{-\lambda_n^2+K\Lambda}\right),\\ &= O\left(n^{-\varepsilon}\right),\end{aligned}$$

uniformly for $|\theta| \leq \varepsilon\sqrt{\Lambda}$.

When $z \in \mathcal{H}_1$, we make the change of variables $z \mapsto 1-\tau/n$ and apply the estimate (50), which gives

$$\begin{aligned}Q_k\left(1-\frac{\tau}{n}, \frac{i\theta}{\sigma_{n,k}}\right) = \frac{n}{\tau}\exp\Bigg\{&\frac{\lambda_n^k}{k!\sigma_{n,k}} i\theta\left(1+O\left(\frac{|\log\tau|}{\Lambda}\right)\right) - \frac{\theta^2}{2}\left(1+O\left(\frac{|\log\tau|}{\Lambda}\right)\right)\\ &+ O\left(\Lambda\sum_{m\geq 3}\frac{|\theta|^m}{m!\Lambda^{m/2}}\right)\Bigg\}.\end{aligned}$$

From this we deduce that if $|\theta| \leq \varepsilon\Lambda^{1/6}$, then

$$Q_k\left(1-\frac{\tau}{n}, \frac{i\theta}{\sigma_{n,k}}\right) = \frac{n}{\tau}\exp\left(\frac{\lambda_n^k}{k!\sigma_{n,k}} i\theta - \frac{\theta^2}{2}\right)\left(1+O\left((|\theta|+|\theta|^3)\frac{|\log\tau|}{\sqrt{\Lambda}}\right)\right);$$

and if $\varepsilon\Lambda^{1/6} \leq |\theta| \leq \varepsilon\Lambda^{1/2}$, then

$$\begin{aligned}Q_k\left(1-\frac{\tau}{n}, \frac{i\theta}{\sigma_{n,k}}\right) &= O\left(\frac{n}{|\tau|}|\tau|^{-\varepsilon}e^{-\theta^2/2+K|\theta|^3/\sqrt{\Lambda}}\right)\\ &= O\left(\frac{n}{|\tau|^{1-\varepsilon}}e^{-\theta^2/4}\right),\end{aligned}$$

for sufficiently small ε.

These estimates then yield

$$\mathbb{E}(e^{X_{n,k}^* i\theta}) = \frac{e^{-\theta^2/2}}{2\pi i}\int_{\mathcal{H}_0}\frac{e^\tau}{\tau}\left(1+O\left((|\theta|+|\theta|^3)\frac{|\log\tau|}{\sqrt{\Lambda}}\right)\right)\left(1+O\left(\frac{|\tau|^2}{n}\right)\right)\mathrm{d}\tau + O\left(n^{-\varepsilon}\right)$$
$$= e^{-\theta^2/2}\left(1+O\left(\frac{|\theta|+|\theta|^3}{\sqrt{\Lambda}}\right)\right)+O\left(n^{-\varepsilon}\right),$$

uniformly for $|\theta| \le \varepsilon\Lambda^{1/6}$, where the contour $\mathcal{H}_0$ is shown in Figure 3, and similarly

$$\mathbb{E}(e^{X_{n,k}^* i\theta}) = O\left(e^{-\theta^2/4}+n^{-\varepsilon}\right),$$

uniformly for $\varepsilon\Lambda^{1/6} \le |\theta| \le \varepsilon\Lambda^{1/2}$. This completes the proof of Proposition 2.

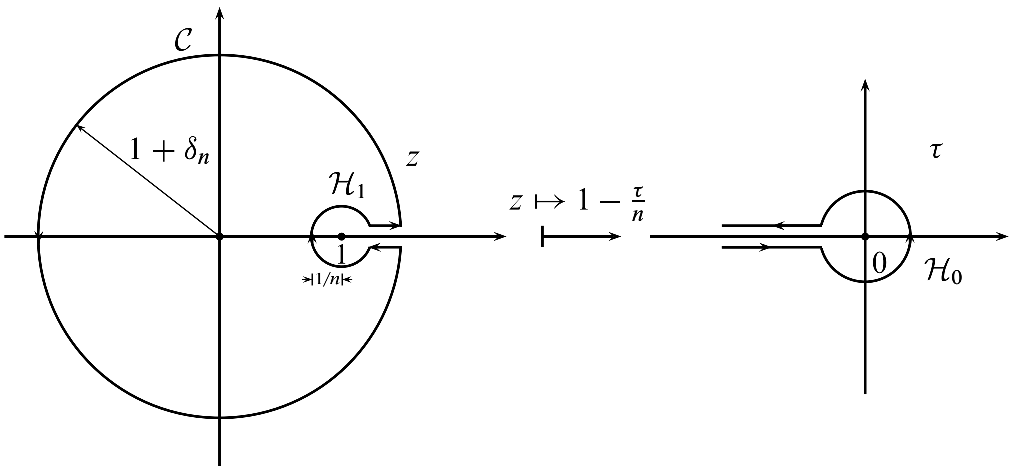

Figure 3: *The Hankel contours used to derive the asymptotics of the moments of $X_{n,k}$.*

Proof of Theorem 3. We now apply the Berry-Esseen smoothing inequality (see Petrov, 1975)

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(X_{n,k}^* < x\right)-\Phi(x)\right| = O\left(\frac{1}{\sqrt{\Lambda}}+J\right),$$

where

$$J = \int_{-\varepsilon\sqrt{\Lambda}}^{\varepsilon\sqrt{\Lambda}}\left|\frac{\mathbb{E}\left(e^{X_{n,k}^* i\theta}\right)-e^{-\theta^2/2}}{\theta}\right|\mathrm{d}\theta$$
$$= \left(\int_{|\theta|\le\Lambda^{-1/2}}+\int_{\Lambda^{-1/2}\le|\theta|\le\varepsilon\Lambda^{1/6}}+\int_{\varepsilon\Lambda^{1/6}\le|\theta|\le\varepsilon\Lambda^{1/2}}\right)\left|\frac{\mathbb{E}\left(e^{X_{n,k}^* i\theta}\right)-e^{-\theta^2/2}}{\theta}\right|\mathrm{d}\theta$$
$$=: J_1+J_2+J_3.$$

The integral J_1 is assessed as follows.

$$\begin{aligned}
J_1 &\le \int_{|\theta|\le\Lambda^{-1/2}} \left|\frac{\mathbb{E}\left(e^{X^*_{n,k}i\theta}\right)-1}{\theta}\right| \mathrm{d}\theta + \int_{|\theta|\le\Lambda^{-1/2}} \left|\frac{e^{-\theta^2/2}-1}{\theta}\right| \mathrm{d}\theta \\
&\le \mathbb{E}(X^{*2}_{n,k}) \int_{|\theta|\le\Lambda^{-1/2}} |\theta|\,\mathrm{d}\theta + \int_{|\theta|\le\Lambda^{-1/2}} |\theta|\,\mathrm{d}\theta \\
&= O(\Lambda^{-1}).
\end{aligned}$$

By (44), the integral J_2 satisfies

$$\begin{aligned}
J_2 &= O\left(\Lambda^{-1/2}\int_{\Lambda^{-1/2}\le|\theta|\le\varepsilon\Lambda^{1/6}} (1+\theta^2)e^{-\theta^2/2}\,\mathrm{d}\theta + n^{-\varepsilon}\int_{\Lambda^{-1/2}\le|\theta|\le\varepsilon\Lambda^{1/6}} |\theta|^{-1}\,\mathrm{d}\theta\right) \\
&= O\left(\Lambda^{-1/2} + n^{-\varepsilon}\log\Lambda\right) \\
&= O\left(\Lambda^{-1/2}\right).
\end{aligned}$$

The last integral J_3 is estimated by using (45)

$$\begin{aligned}
J_3 &= O\left(\int_{\varepsilon\Lambda^{1/6}\le|\theta|\le\varepsilon\Lambda^{1/2}} \theta^{-1}e^{-\theta^2/4}\,\mathrm{d}\theta + n^{-\varepsilon}\log\Lambda\right) \\
&= O\left(\Lambda^{-1/2}\right).
\end{aligned}$$

This proves Theorem 3. ∎

In particular, Theorem 3 implies and completes the case $\alpha = 0$ in Theorem 1.

An alternative proof of Theorem 1 (*ii*). The above approach based on VW-polynomials can also be refined to give an alternative proof of Theorem 1. We outline the main steps.

First, by (47) and induction, we can prove that

$$\begin{cases}
V_{k,m}(x) \sim \xi_m\left(\dfrac{k}{x}\right)\dfrac{(k/x)^{m-1}}{m}\cdot\dfrac{x^{mk}}{k!^m}, \\
W_{k,m}(x) \sim \xi_m\left(\dfrac{k}{x}\right)\dfrac{x^{mk}}{k!^m},
\end{cases}$$

uniformly for $0 < k/|x| < m^{1/(m-1)}$ and large complex x, where $\xi_m(u)$ is defined recursively by

$$\xi_m(u) = \frac{1}{m-u^{m-1}} \sum_{1\le h<m} \binom{m}{h} \xi_h(u)\xi_{m-h}(u)u^{h-1} \qquad (m \ge 2),$$

with $\xi_1(u) = 1$.

Then when $k/\lambda_n \to \alpha$, $0 < \alpha < m^{1/(m-1)}$,

$$\begin{aligned}\mathbb{E}(X_{n,k}^m) &= [z^n]\frac{z}{1-z}W_{k,m}(L(z))\\ &\sim \frac{1}{2\pi i}\int_{\mathcal{H}} e^{\tau}\tau^{-1}W_{k,m}(\log(n/\tau))\,\mathrm{d}\tau\\ &\sim \frac{\xi_m(\alpha)}{2\pi i}\int_{\mathcal{H}} e^{\tau}\tau^{-1}\frac{(\lambda_n - \log\tau)^{mk}}{k!^m}\,\mathrm{d}\tau\\ &\sim \xi_m(\alpha)\frac{\lambda_n^{mk}}{k!^m}\frac{1}{2\pi i}\int_{\mathcal{H}} e^{\tau}\tau^{-1-m\alpha}\,\mathrm{d}\tau\\ &\sim \frac{\xi_m(\alpha)}{\Gamma(1+m\alpha)}\cdot\frac{\lambda_n^{mk}}{k!^m}\\ &\sim \xi_m(\alpha)\frac{\Gamma(1+\alpha)^m}{\Gamma(1+m\alpha)}\mu_{n,k}^m,\end{aligned}$$

for a suitably chosen Hankel contour $\mathcal{H}$. And it is straightforward to check, by (7), that

$$\xi_m(\alpha)\frac{\Gamma(1+\alpha)^m}{\Gamma(1+m\alpha)} = \nu_m(\alpha).$$

Note that this approach does not apply to profiles of binary search trees.

7 Profiles of random binary search trees

We consider briefly in this section random binary search trees whose profiles have been widely studied; see Drmota and Hwang (2005a) and the references therein. Our method of moments and contraction method apply. While the results for both trees are very similar, there is no range for binary search trees where the limit law of the profile is normal.

Let $Y_{n,k}$ denote the number of external nodes at distance k from the root and $Z_{n,k}$ the number of internal nodes at level k (root being at level 0) in a random binary search tree of n nodes (as constructed from a random permutation of n elements). Then for $k, n \geq 1$

$$\begin{aligned}Y_{n,k} &\overset{\mathscr{D}}{=} Y_{J_n,k-1} + Y^*_{n-1-J_n,k-1},\\ Z_{n,k} &\overset{\mathscr{D}}{=} Z_{J_n,k-1} + Z^*_{n-1-J_n,k-1},\end{aligned}$$

with the initial conditions $Y_{n,0} = \delta_{n,0}$ and $Z_{n,0} = 1 - \delta_{n,0}$, where J_n is uniformly distributed over $\{0,\ldots,n-1\}$, the summands are independent and $Y_{n,k} \overset{\mathscr{D}}{=} Y^*_{n,k}$, $Z_{n,k} \overset{\mathscr{D}}{=} Z^*_{n,k}$. Note that $Z_{n,k} = \sum_{j>k} Y_{n,j}2^{j-k}$.

Mean values. The expected value of $Y_{n,k}$ satisfies (see Drmota and Hwang, 2005a and the references therein)

$$\mathbb{E}(Y_{n,k}) = \frac{2^k}{n!}\mathsf{s}(n,k) = \frac{(2\lambda_n)^k}{\Gamma(\alpha_{n,k})k!n}\left(1 + O\left(\frac{1}{\lambda_n}\right)\right),$$

the O-term holding uniformly for $1 \leq k \leq K\lambda_n$.

For internal nodes, the asymptotic behavior is different

$$
\begin{aligned}
\mathbb{E}(Z_{n,k}) &= \frac{2^k}{n!}\sum_{j>k} \mathsf{s}(n,j) \\
&\sim \begin{cases} 2^k - \dfrac{(2\lambda_n)^k}{(1-\alpha_{n,k})\Gamma(\alpha_{n,k})nk!}, & \text{if } 1 \le k \le \lambda_n - K\sqrt{\lambda_n}; \\ 2^k \Phi(-x_{n,k}), & \text{if } x_{n,k} := (k-\lambda_n)/\sqrt{\lambda_n} = o((\lambda_n)^{1/6}); \\ \dfrac{(2\lambda_n)^k}{(\alpha_{n,k}-1)\Gamma(\alpha_{n,k})nk!}, & \text{if } \lambda_n + K\sqrt{\lambda_n} \le k \le K\lambda_n, \end{cases}
\end{aligned}
$$

where the error terms in the first and the third approximations are of the form

$$
O\left(\frac{(2\lambda_n)^k}{|k-\lambda_n|^2 nk!}\right),
$$

and that of the middle is $O((1+|x_{n,k}|^3)/\sqrt{\lambda_n})$; see (51) below.

Note that

$$
\frac{\log \mathbb{E}(Y_{n,k})}{\lambda_n} \to \alpha - 1 - \alpha\log(\alpha/2),
$$

and the right-hand side is positive when $\alpha_- < \alpha < \alpha_+$, where $0 < \alpha_- < 1 < \alpha_+$ are the two real zeros of the equation $z - 1 - z\log(z/2)$ or $e^{(z-1)/z} = z/2$. These two constants are sometimes referred to as the *binary search tree constants* (or the fill-up level and height constants, respectively).

The limit law. Define the map

$$
T : \mathcal{M} \to \mathcal{M}, \quad \nu \mapsto \mathcal{L}\left(\frac{\alpha}{2}U^{\alpha-1}Z + \frac{\alpha}{2}(1-U)^{\alpha-1}Z^*\right),
$$

where Z, Z^*, U are independent and $\mathcal{L}(Z) = \mathcal{L}(Z^*) = \nu$.

The constant s is defined by $s := 2$ when $2-\sqrt{2} < \alpha < 2+\sqrt{2}$ and $1 < s < \varrho$ when $\alpha \in (\alpha_-, \alpha_+) \setminus (2-\sqrt{2}, 2+\sqrt{2})$, where $\varrho \in (1,2]$ solves the equation $\varrho(\alpha-1)+1 = 2(\alpha/2)^\varrho$.

Similar to Proposition 1, we have the following properties.

Proposition 3. *If $\alpha_- < \alpha < \alpha_+$, then the restriction of T to $\mathcal{M}_s(1)$ has a unique fixed point $Y(\alpha)$. In addition, $\mathbb{E}|Y(\alpha)|^\varrho = \infty$ for $\alpha \in (\alpha_-, \alpha_+) \setminus (2-\sqrt{2}, 2+\sqrt{2})$.*

Limit distribution when $\alpha_- < \alpha < \alpha_+$. The above estimates for the mean values of $Y_{n,k}$ and $Z_{n,k}$ say roughly that internal nodes are asymptotically full (of sizes 2^k) for the first $\lambda_n - K\sqrt{\lambda_n}$ levels, while external nodes are relatively sparse there. Observe that the second order term of $\mathbb{E}(Z_{n,k})$ is asymptotically of the same order as $\mathbb{E}(Y_{n,k})$ when $\alpha < 1$. This suggests that we should consider

$$
\bar{Z}_{n,k} := \begin{cases} 2^k - Z_{n,k}, & \text{if } \alpha_- \le \alpha < 1; \\ Z_{n,k}, & \text{if } 1 \le \alpha < \alpha_+. \end{cases}
$$

Theorem 5. *Let $Y(\alpha)$ and ϱ be defined as in Proposition 3. Assume that $k = \alpha\lambda_n + o(\lambda_n)$. Then for $\alpha_- < \alpha < \alpha_+$,*

$$
\frac{Y_{n,k}}{\mathbb{E}(Y_{n,k})}, \frac{\bar{Z}_{n,k}}{\mathbb{E}(\bar{Z}_{n,k})} \xrightarrow{\mathscr{D}} Y(\alpha),
$$

with convergence of all moments for $\alpha \in [1,2]$ but not for α outside $[1,2]$.

Chauvin et al. (2005) proved almost sure convergence for $Y_{n,k}/\mathbb{E}(Y_{n,k})$ when $\alpha_- < \alpha < \alpha_+$; their result is stronger than convergence in distribution but does not imply convergence of all moments.

As in Theorem 4, we can derive a convergence rate for the ζ_2-distance when $2-\sqrt{2} < \alpha < 2+\sqrt{2}$ and for ζ_s when $\alpha \in (\alpha_,\alpha_+) \setminus (2-\sqrt{2}, 2+\sqrt{2})$.

Moments of the limit law. The integral moments $\eta_m(\alpha)$ of $Y(\alpha)$ satisfy (when they exist) $\eta_0(\alpha) = \eta_1(\alpha) = 1$ and for $m \ge 2$

$$\eta_m(\alpha) = \frac{(\alpha/2)^m}{m(\alpha-1)+1-2(\alpha/2)^m} \sum_{1\le h<m} \binom{m}{h} \eta_h(\alpha)\eta_{m-h}(\alpha) \frac{\Gamma(h(\alpha-1)+1)\Gamma((m-h)(\alpha-1)+1)}{\Gamma(m(\alpha-1)+1)}.$$

Observe that the polynomial $m(z-1)+1-2(z/2)^m$ has two positive zeros z_m^- and z_m^+, where $z_m^- \in [2-\sqrt{2}, 1)$ and $z_m^+ \in (2, 2+\sqrt{2}]$ for $m \ge 2$. And the two sequences of zeros for increasing m satisfy (see Table 1)

$$z_m^- \uparrow 1, \qquad z_m^+ \downarrow 2.$$

Thus the interval $[1,2]$ is the only range where convergence of all moments holds.

More precisely, $\eta_m(\alpha)$ is finite when $z_m^- < \alpha < z_m^+$ and we have convergence of the first m-th moment (but not the $(m+1)$-st moment) for $Y_{n,k}/\mathbb{E}(Y_{n,k})$ and $\bar{Z}_{n,k}/\mathbb{E}(\bar{Z}_{n,k})$ there. In particular, if $\alpha_- < \alpha \le 2-\sqrt{2}$ or $2+\sqrt{2} \le \alpha < \alpha_+$, then $Y(\alpha)$ has no second moment. This is consistent with the result in Drmota and Hwang (2005a).

m	2	3	4	5	6
z_m^-	0.58578	0.69459	0.76045	0.80420	0.83509
z_m^+	3.41421	3.06417	2.86989	2.74376	2.65416
m	7	8	9	10	11
z_m^-	0.85790	0.87533	0.88903	0.90006	0.90912
z_m^+	2.58668	2.53372	2.49085	2.45532	2.42531

Table 1: *Approximate numeric values of z_m^- and z_m^+ for $m = 2, \ldots, 11$.*

Limit distributions when $\alpha = 1$. Note that $Y(1) = Y(2) \equiv 1$.

The following theorem states that there is a delicate difference between the limit distribution of $Y_{n,k}$ and that of $Z_{n,k}$ (properly normalized) when $\alpha = 1 + O(1/\sqrt{\lambda_n})$.

Theorem 6. *Assume $k = \lambda_n + t_{n,k}$, where $t_{n,k} = o(\lambda_n)$. If $|t_{n,k}| \to \infty$, then*

$$\frac{Y_{n,k} - \mathbb{E}(Y_{n,k})}{2t_{n,k}(2\lambda_n)^{k-1}/(nk!)} \xrightarrow{\mathscr{M}} Y'(1);$$

if $t_{n,k} = O(1)$, then the sequence of random variables $(Y_{n,k} - \mathbb{E}(Y_{n,k}))/\sqrt{\mathbb{V}(Y_{n,k})}$ does not converge to a fixed limit law.

For internal nodes, uniformly for $t_{n,k} = o(\lambda_n)$,

$$\frac{Z_{n,k} - \mathbb{E}(Z_{n,k})}{(2\lambda_n)^k/(nk!)} \xrightarrow{\mathscr{M}} Y'(1).$$

Thus the periodicity does not play a special role for internal nodes when $\alpha = 1$. Note that the normalizing constants differ by the factor $\alpha_{n,k} - 1 = t_{n,k}/\lambda_n$.

The limit law $Y'(1)$ can also be defined as

$$Y'(1) \overset{\mathscr{D}}{=} \frac{1}{2}Y'(1) + \frac{1}{2}Y'(1)^* + 1 + \frac{1}{2}\log U + \frac{1}{2}\log(1-U),$$

with independent summands and $Y'(1) \overset{\mathscr{D}}{=} Y'(1)^*$. Note that the random variables $\sum_{j\ge 0} Z_{n,j}/2^j$ have mean equal to $\sum_{1\le j\le n} j^{-1}$ and converge to $Y'(1)$ (after centered and normalized).

Since the distribution of $Y'(1)$ is uniquely characterized by its moment sequence, the convergence in distribution is also implied by the Frechet-Shohat moment convergence theorem.

The quicksort limit law when $\alpha = 2$.

Theorem 7. *Assume $\alpha_{n,k} = 2 + t_{n,k}/\lambda_n$, where $t_{n,k} = o(\lambda_n)$. If $|t_{n,k}| \to \infty$, then*

$$\frac{Y_{n,k} - \mathbb{E}(Y_{n,k})}{2t_{n,k}(2\lambda_n)^{k-1}/(nk!)}, \ \frac{Z_{n,k} - \mathbb{E}(Z_{n,k})}{2t_{n,k}(2\lambda_n)^{k-1}/(nk!)} \xrightarrow{\mathscr{M}} Y'(2);$$

if $t_{n,k} = O(1)$, then neither of two sequences

$$\left\{ \frac{Y_{n,k} - \mathbb{E}(Y_{n,k})}{\sqrt{\mathbb{V}(Y_{n,k})}}, \frac{Z_{n,k} - \mathbb{E}(Z_{n,k})}{\sqrt{\mathbb{V}(Z_{n,k})}} \right\}$$

converges to a fixed limit law.

The limit law $Y'(2)$ is essentially the quicksort limit law (see Hwang and Neininger, 2002)

$$Y'(2) \overset{\mathscr{D}}{=} UY'(2) + (1-U)Y'(2)^* + \frac{1}{2} + U\log U + (1-U)\log(1-U),$$

with independent summands on the right-hand side and $Y'(2) \overset{\mathscr{D}}{=} Y'(2)^*$.

Convergence in distribution in the case when $|t_{n,k}| \to \infty$ is also implied.

The approach given in this paper gives not only the bimodality of the variances $\mathbb{V}(Y_{n,k})$ and $\mathbb{V}(Z_{n,k})$ but also the extremal (reachable) orders of $|\mathbb{E}(Y_{n,k} - \mathbb{E}(Y_{n,k}))^m|$ and $|\mathbb{E}(Z_{n,k} - \mathbb{E}(Z_{n,k}))^m|$ for $m \ge 3$ when $\alpha = 2$.

Sketch of proofs. We sketch a few steps for internal nodes, external nodes being similar and simpler.

Starting from the recurrence for the probability generating function of $Z_{n,k}$

$$P_{n,k}(y) = \frac{1}{n}\sum_{0\le j<n} P_{j,k-1}(y)P_{n-1-j,k-1}(y) \qquad (n \ge 2; k \ge 1),$$

with $P_{0,0}(y) = 1$ and $P_{n,0}(y) = y$ for $n \ge 1$, we have the recurrence for the mean value

$$\mathbb{E}(Z_{n,k}) = \frac{2}{n}\sum_{0\le j<n} \mathbb{E}(Z_{j,k-1}) \qquad (n \ge 2; k \ge 1).$$

Lemma 9. *The solution to the recurrence*

$$a_{n,k} = \frac{2}{n} \sum_{0 \le j < n} a_{j,k-1} + b_{n,k},$$

is given explicitly by

$$a_{n,k} = b_{n,k} + \frac{2}{n} \sum_{0 \le j < n} \sum_{0 \le r < k} b_{j,k-1-r} [u^r] \prod_{j < \ell < n} \left(1 + \frac{2u}{\ell}\right),$$

where $b_{0,k} := a_{0,k}$.

Then we have, by applying the exact solution with $b_{n,0} = 1$ for $n \ge 1$ and $b_{n,k} = 0$ otherwise,

$$\begin{aligned}
\mathbb{E}(Z_{n,k}) &= \frac{2}{n}[u^{k-1}] \sum_{1 \le j < n} \prod_{j < \ell < n} \left(1 + \frac{2u}{\ell}\right) \\
&= 2^k [u^{k-1}] \frac{1}{u-1} \left(\frac{\Gamma(n+u)}{\Gamma(n+1)\Gamma(u+1)} - 1 \right) \\
&= \frac{2^k}{2\pi i} \oint_{|u|=\alpha_{n,k}>1} u^{-k-1} \frac{1}{u-1} \binom{n+u-1}{n} \, \mathrm{d}u. \qquad (51)
\end{aligned}$$

Thus

$$\begin{aligned}
&\mathbb{E}(Z_{j,k-1}) + \mathbb{E}(Z_{n-1-j,k-1}) - \mathbb{E}(Z_{n,k}) \\
&\quad = \frac{2^k}{2\pi i} \oint_{|u|=\alpha_{n,k}} u^{-k-1} n^{u-1} \phi(u, j/n) \left(1 + O(j^{-1} + (n-j)^{-1})\right) \mathrm{d}u,
\end{aligned}$$

where

$$\phi(u, x) = \frac{u x^{u-1} + u(1-x)^{u-1} - 2}{2\Gamma(u)(u-1)}.$$

Note that, unlike recursive trees and external nodes of binary search trees, $\phi(1, x)$ is not zero and $\phi(1, x) = 1 + \frac{1}{2}\log x + \frac{1}{2}\log(1-x)$. This is why there is no periodic case for internal nodes when $\alpha = 1 + O(1/\sqrt{\lambda_n})$.

All estimates required for $\mathbb{E}(Z_{n,k})$ and for its difference $\mathbb{E}(Z_{j,k-1}) + \mathbb{E}(Z_{n-1-j,k-1}) - \mathbb{E}(Z_{n,k})$ can be derived as for recursive trees. For example, we have, uniformly for $\lambda_n + K\sqrt{\lambda_n} \le k \le K\lambda_n$,

$$\mathbb{E}(Z_{n,k}) \sim \frac{(2\lambda_n)^k}{(\alpha - 1)\Gamma(\alpha) k! n}.$$

8 Conclusions

Most random trees in discrete probability or data structures have height of order either in $\sqrt{n}$ or in $\log n$; see Aldous (1991). While profiles and other related processes defined on random trees of $\sqrt{n}$-height have been thoroughly studied in the literature (see Aldous, 1991, Drmota and Gittenberger, 1997, Kersting, 1998, Pitman, 1999, and the references therein), profiles of trees with logarithmic height have received little attention (except for digital search trees; see Aldous and Shields, 1988, Jacquet et al., 2001). This paper shows that the phenomena exhibited in such trees are drastically different yet highly attractive.

A detailed study of more general random search trees (including m-ary search trees, quadtrees, fringe-balanced binary search trees, etc.) will be given elsewhere.

Many questions remain unclear at this stage. For example, are there more "humps" or valleys for higher central moments or cumulants in the central range? Are there interesting process approximations? How to simulate the limit laws appearing in this paper? And what happens when $\alpha = e$ for recursive trees and $\alpha = \alpha_-, \alpha_+$ for binary search trees? Do we still have the same convergence in distribution for $X_{n,k}/\mu_{n,k}$ when $\mu_{n,k} \to \infty$? Note that for recursive trees, $\mathbb{E}(X_{n,k}) \to \infty$ for $k \le e\lambda_n - e_1 \log \lambda_n$, where $e_1 > 1/2$, but $\mathbb{V}(X_{n,k}) \to \infty$ for $k \le \frac{4}{\log 4}\lambda_n - e_2 \log \lambda_n$, where $e_2 > 1/(2\log 4)$. Since $4/\log 4 \approx 2.88 > e$, there is still a small range in k where the mean goes to zero but the variance goes to infinity.

Acknowledgements

Part of the work of the second author was carried out while he was visiting Institut für Stochastik und Mathematische Informatik, J. W. Goethe-Universität (Frankfurt) and Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien; he thanks both Institutes for their hospitality and support.

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] D. Aldous and P. Shields (1988). A diffusion limit for a class of randomly-growing binary trees. *Probability Theory and Related Fields* **79** 509–542.

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

[4] D. Y. C. Chan, B. D. Hughes, A. S. Leong and W. J. Reed (2003). Stochastically evolving networks. *Physical Review E* **68** 06624-1–06624-24.

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

[6] B. Chauvin, T. Klein, J.-F. Marckert and A. Rouault (2005). Martingales, embedding and tilting of binary trees. *Electronic Journal of Probability* to appear.

[7] B. Chauvin and A. Rouault (2004). Connecting Yule process, bisection and binary search tree via martingales. *Journal of the Iranian Statistical Society 3* 88–116.

[8] 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.

[9] 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.

[10] H.-H. Chern, H.-K. Hwang and T.-H. Tsai (2002). An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms. *Journal of Algorithms* **44** 177–225.

[11] L. Devroye (1987). Branching processes in the analysis of the heights of trees. *Acta Informatica* **24** 277–298.

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

[13] L. Devroye and J. M. Robson (1995). On the generation of random binary search trees. *SIAM Journal on Computing* **24** 1141–1156.

[14] L. Devroye, C. McDiarmid and B. Reed (2002). Giant components for two expanding graph processes. In *Mathematics and Computer Science, II (Versailles, 2002)*, 161–173, Birkhäuser, Basel, 2002.

[15] R. P. Dobrow and J. A. Fill (1999). Total path length for random recursive trees. *Combinatorics, Probability and Computing* **8** 317–333.

[16] M. Dondajewski and J. Szymański (1982). On the distribution of vertex-degrees in a strata of a random recursive tree. *Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques* **30** 205–209.

[17] M. Drmota and B. Gittenberger (1997). On the profile of random trees. *Random Structures and Algorithms* **10** 421–451.

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

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

[20] P. Erdős and P. Turán (1967). On some problems of a statistical group-theory. III. *Acta Mathematica Academiae Scientiarum Hungaricae* **18** 309–320.

[21] D. J. Evans and R. C. Dunbar (1982). The parallel quicksort algorithm. I. Run time analysis. *International Journal of Computer Mathematics* **12** 19–55.

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

[23] J. L. Gastwirth and P. K. Bhattacharya (1984). Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable. *Operations Research* **32** 527–536.

[24] R. Grossman and R. G. Larson (1989). Hopf-algebraic structure of families of trees. *Journal of Algebra* **126** 184–210.

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

[26] H.-K. Hwang (2005). Profiles of random trees: plane-oriented recursive trees. Preprint submitted for publication.

[27] H.-K. Hwang and R. Neininger (2002). Phase change of limit laws in the quicksort recurrence under varying toll functions. *SIAM Journal on Computing* **31** 1687–1722.

[28] P. Jacquet, W. Szpankowski and J. Tang (2001). Average profile of the Lempel-Ziv parsing scheme for a Markovian source. *Algorithmica* **31** 318–360.

[29] M. Janic, F. Kuipers, X. Zhou and P. van Mieghem (2002). Implications for QoS provisioning based on traceroute measurements. In *Lecture Notes in Computer Science 2511*, pp 3–14.

[30] G. Kersting, On the height profile of a conditioned Galton-Watson tree. Preprint (1998).

[31] D. E. Knuth (1997) *The Art of Computer Programming, Volume 1, Fundamental Algorithms*. Second Edition, Addison-Wesley, Reading, MA.

[32] H. M. Mahmoud (1991). Limiting distributions for path lengths in recursive trees. *Probability in the Engineering and Informational Sciences* 5 53–59.

[33] H. M. Mahmoud (1992). *Evolution of Random Search Trees*. John Wiley & Sons, New York.

[34] H. M. Mahmoud and R. T. Smythe (1995). Probabilistic analysis of bucket recursive trees. *Theoretical Computer Science* **144** 221–249.

[35] A. Meir and J. W. Moon (1974). Cutting down recursive trees. *Mathematical Biosciences* **21** 173–181.

[36] A. Meir and J. W. Moon (1978). On the altitude of nodes in random trees. *Canadian Journal of Mathematics* **30** 997–1015.

[37] J. W. Moon (1974). The distance between nodes in recursive trees. In "*Combinatorics*", edited by T. P. McDonough and V. C. Marron. London Mathematical Society Lecture Notes, Series 13, London, pp. 125–132.

[38] H. S. Na and A. Rapoport (1970). Distribution of nodes of a tree by degree. *Mathematical Biosciences* **6** 313–329.

[39] D. Najock and C. C. Heyde (1982). On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *Journal of Applied Probability* **19** 675–680.

[40] R. Neininger and L. Rüschendorf (2004). A general limit theorem for recursive algorithms and combinatorial structures. *Annals of Applied Probability* **14** 378-418.

[41] V. V. Petrov (1975). *Sums of Independent Random Variables.* Springer-Verlag, Heidelberg; translated from the Russian by A. A. Brown.

[42] J. Pitman (1999). The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree or forest. *Annals of Probability* **27** 261–283.

[43] S. T. Rachev (1991). *Probability Metrics and the Stability of Stochastic Models*. John Wiley & Sons Ltd., Chichester.

[44] U. Rösler and L. Rüschendorf (2001). The contraction method for recursive algorithms. *Algorithmica* **29** 3–33.

[45] R. T. Smythe and H. M. Mahmoud (1995). A survey of recursive trees. *Theory of Probability and Mathematical Statistics* **51** 1–27.

[46] J. Szymański (1990). On the maximum degree and the height of a random recursive tree. In *Random Graphs* '87 (Poznań, 1987), 313–324, Wiley, Chichester, 1990.

[47] G. T. Tetzlaff (2002). Breakage and restoration in recusrive trees. *Journal of Applied Probability* **39** 383–390.

[48] P. van Mieghem, G. Hooghiemstra and R. van der Hofstad (2001). On the efficiency of multicast. *IEEE/ACM Transactions on Networking* **9** 719–732.

[49] R. van der Hofstad, G. Hooghiemstra and P. van Mieghem (2002). On the covariance of the level sizes in random recursive trees. *Random Structures and Algorithms* **20** 519–539.

[50] V. M. Zolotarev (1976). Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces. *Theory of Probability and its Applications* **21** 721–737.

[51] V. M. Zolotarev (1977). Ideal metrics in the problem of approximating distributions of sums of independent random variables. *Theory of Probability and its Applications* **22** 433–449.