THE CRITICAL BETA-SPLITTING RANDOM TREE IV: MELLIN ANALYSIS OF LEAF HEIGHT

DAVID J. ALDOUS AND SVANTE JANSON

Abstract. In the critical beta-splitting model of a random n-leaf rooted tree, clades are recursively split into sub-clades, and a clade of m leaves is split into sub-clades containing i and $m-i$ leaves with probabilities $\propto 1/(i(m-i))$. The height of a uniform random leaf can be represented as the absorption time of a certain *harmonic descent* Markov chain. Recent work on these heights D_n and L_n (corresponding to discrete or continuous versions of the tree) has led to quite sharp expressions for their asymptotic distributions, based on their Markov chain description. This article gives even sharper expressions, based on an $n \to \infty$ limit tree structure described via exchangeable random partitions in the style of Haas et al (2008). Within this structure, calculations of moments lead to expressions for Mellin transforms, and then via Mellin inversion we obtain sharp estimates for the expectation, variance, Normal approximation and large deviation behavior of D_n.

Keywords: `Exchangeable partition, Markov chain, Mellin transform, random tree, subordinator.`

MSC 60C05; 05C05, 44A15, 60G09.

Contents

Date: December 14, 2024.

SJ supported by the Knut and Alice Wallenberg Foundation and the Swedish Research Council.

1. Introduction

1.1. **The tree model.** A more detailed account of the model, with graphics, is given in [3], and this article is in some ways a continuation of that article.

For $m \geq 2$, consider the probability distribution $(q(m,i),\ 1 \leq i \leq m-1)$ constructed to be proportional to $\frac{1}{i(m-i)}$. Explicitly (by writing $\frac{1}{i(m-i)} = \left(\frac{1}{i} + \frac{1}{m-i}\right)/m$)

$$q(m,i) = \frac{m}{2h_{m-1}} \cdot \frac{1}{i(m-i)},\ 1 \leq i \leq m-1, \tag{1.1}$$

where h_{m-1} is the harmonic sum $\sum_{i=1}^{m-1} 1/i$.

Now fix $n \geq 2$. Consider the process of constructing a random tree by recursively splitting the integer interval $[n] = \{1, 2, \ldots, n\}$ of "leaves" as follows. First specify that there is a left edge and a right edge at the root, leading to a left subtree which will have the[1] G_n leaves $\{1, \ldots, G_n\}$ and a right subtree which will have the remaining $R_n = n - G_n$ leaves $\{G_n + 1, \ldots, n\}$, where G_n (and also R_n, by symmetry) has distribution $q(n, \cdot)$. Recursively, a subinterval with $m \geq 2$ leaves is split into two subintervals of random size from the distribution $q(m, \cdot)$. Continue until reaching intervals of size 1, which are the leaves. This process has a natural tree structure. In this discrete-time construction, which we call DTCS(n), we regard the edges of the tree as having length 1. It turns out (see e.g. [2]) to be natural to consider also the continuous-time construction CTCS(n) in which a size-m interval

[1] G for *gauche* (left) because we use L_n for leaf height.

is split at rate h_{m-1}, that is after an Exponential(h_{m-1}) holding time. Once such a tree is constructed, it is natural to identify "time" with "distance": a leaf that appears at time t has *height* t. Of course the discrete-time model is implicit within the continuous-time model, and a leaf in DTCS(n) which appears after ℓ splits has *hop-height* ℓ.

Finally, our results do not use the leaf-labels $\{1, 2, \ldots, n\}$ in the interval-splitting construction. Instead they involve a uniform *random* leaf. Equivalently, one could take a uniform random permutation of labels and then talk about the leaf with some arbitrary label.

Many aspects of this model can be studied by different techniques, see [2] for an overview.In this article we focus on one aspect and one methodology, as follows.

1.2. **Leaf heights.** It is an elementary calculation [5] to show that the discrete time process described by

> In the path from the root to a uniform random leaf in DTCS(n), consider at each step the size (number of leaves) of the sub-tree rooted at the current position

is the discrete time Markov chain $(X_t^{disc}, t = 0, 1, 2, \ldots)$ on states $\{1, 2, 3, \ldots\}$ with transition probabilities

$$q^*_{m,i} := \tfrac{1}{h_{m-1}} \tfrac{1}{m-i}, \ 1 \leq i \leq m-1, m \geq 2. \tag{1.2}$$

Similarly, the continuous time process described by

> Move at speed one along the edges of the path from the root to a uniform random leaf of CTCS(n), and consider at each time the size (number of leaves) of the sub-tree rooted at the current position

is the continuous time Markov chain $(X_t^{cont}, t \geq 0)$ with transition rates

$$\lambda_{m,i} := \tfrac{1}{m-i}, \ 1 \leq i \leq m-1, m \geq 2. \tag{1.3}$$

Each process is absorbed at state 1. So if we define

$D_n :=$ hop-height of a uniform random leaf of DTCS(n)

$L_n :=$ height of a uniform random leaf of CTCS(n)

then these are the same as the appropriate Markov chain absorption time

$$D_n = \inf\{t : X_t^{cont} = 1 \mid X_0^{cont} = n\} \tag{1.4}$$

$$L_n = \min\{t : X_t^{disc} = 1 \mid X_0^{disc} = n\}. \tag{1.5}$$

We call these Markov chains the *harmonic descent (HD)* chains. Recent studies of D_n and L_n [4, 5, 16, 15, 18] have been based on the Markov chain representation (1.4, 1.5). Note that to study $n \to \infty$ asymptotics for these chains, one cannot directly formalize the idea of starting a hypothetical version of the chain from $+\infty$ at time $-\infty$. However in the underlying random tree model CTCS(n) there is a more sophisticated way to formalize a limit tree structure CTCS(∞) via exchangeable random partitions (Section 2). In the context of our model, this methodology was first used in [3], and this continuation article will show that one can then apply Mellin transform techniques to obtain asymptotic estimates that are sharper than those obtained by previous methods. Here are summaries of our main results.

1.3. **Summary of results.** Let ψ be the digamma function (see Section 3.2). Let $0 > s_1 > s_2 > \dots$ be the negative roots of $\psi(s) = \psi(1)$. Recall that $\zeta(2) = \pi^2/6$ and $\zeta(3) \doteq 1.202$, and note that $\sim$ in the results below denotes asymptotic expansion (see Section 3.1).

Theorem 1.1. *As $n \to \infty$*

$$\mathbb{E}\,[D_n] \sim \frac{6}{\pi^2}\log n + \sum_{i=0}^{\infty} c_i n^{-i} + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} c_{j,k}\, n^{-|s_j|-k} \tag{1.6}$$

for some coefficients c_i and $c_{j,k}$ that can be found explicitly; in particular, $c_1 = -3/\pi^2$ and

$$c_0 = \frac{\zeta(3)}{\zeta(2)^2} + \frac{\gamma}{\zeta(2)} \doteq 0.795155660439. \tag{1.7}$$

Theorem 1.1 is proved in Sections 7.2 and 10.1. It improves on [5, Theorem 1.2 and Proposition 2.3] which gave the initial terms $\frac{6}{\pi^2}\log n + c_0 + c_1 n^{-1}$ with the explicit formula for c_1 but not the formula[2] for c_0. The discussion of the *h-ansatz* in [5] assumes that only integer powers of $1/n$ should appear in the expansion (1.6), but in fact (surprisingly?) the spectrum of powers of n appearing is $\{-i : i \geq 0\} \cup \{-(|s_j| + k) : j \geq 1, k \geq 1\}$.

Note that there is a simple recurrence for $\mathbb{E}\,[D_n]$. The theme of [5] was to exploit "the recurrence method", that is to take a sequence defined by a recurrence and then upper and lower bound the unknown sequence by known sequences. This method was used in [5] for many of the problems in this paper, as indicated in the references below.

Theorem 1.2. *As $n \to \infty$*

$$\begin{aligned}\mathbb{E}\,[L_n] \sim \frac{3}{\pi^2}\log^2 n &+ \Big(\frac{\zeta(3)}{\zeta(2)^2} + \frac{\gamma}{\zeta(2)}\Big)\log n + b_0 \\ &+ \sum_{k=1}^{\infty} a_k n^{-k}\log n + \sum_{k=1}^{\infty} b_k n^{-k} + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} c_{j,k} n^{-|s_j|-k}\end{aligned} \tag{1.8}$$

for some computable constants a_k, b_k, $c_{j,k}$; in particular,

$$b_0 = \frac{3\gamma^2}{\pi^2} + \frac{\zeta(3)}{\zeta(2)^2}\gamma + \frac{\zeta(3)^2}{\zeta(2)^3} + \frac{1}{10} \doteq 0.78234. \tag{1.9}$$

Theorem 1.2 is proved in Sections 8 and 10.2. The first term $\frac{3}{\pi^2}\log^2 n$ was observed long ago in [1]. Using the recurrence method, the coefficient for $\log n$ was found in [5, Theorem 1.2]; that coefficient equals the constant term c_0 in the asymptotic expansion (1.6) of $\mathbb{E}\,[D_n]$.

Theorem 1.3. *As $n \to \infty$*

$$\operatorname{var}[D_n] = \frac{2\zeta(3)}{\zeta(2)^3}\log n + \frac{2\zeta(3)}{\zeta(2)^3}\gamma + \frac{5\zeta(3)^2}{\zeta(2)^4} - \frac{18}{5\pi^2} + O\Big(\frac{\log n}{n}\Big). \tag{1.10}$$

[2] Before knowing the exact value of c_0, numerics gave an estimate that agrees with (1.7) to 10 places.

Theorem 1.3 is proved in Section 11.1. The leading term $\frac{2\zeta(3)}{\zeta(2)^3}\log n$ was found by in [5, Theorem 1.1] by the recursion method. Higher moments of D_n are discussed in Section 11.1.

Theorem 1.4. *For $-\infty < z < 1$ there is a unique real number $\rho(z)$ in $(-1,\infty)$ satisfying $\psi\big(1+\rho(z)\big)-\psi(1)=z$. Then*

$$\mathbb{E}\left[e^{zD_n}\right] = \frac{-z\Gamma(-\rho(z))}{\psi'(1+\rho(z))}\frac{\Gamma(n)}{\Gamma(n-\rho(z))} + O\big(n^{-\sigma_*}\big) \tag{1.11}$$

and

$$\mathbb{E}\left[e^{zD_n}\right] = \frac{-z\Gamma(-\rho(z))}{\psi'(1+\rho(z))}n^{\rho(z)}\cdot\big(1+O\big(n^{-\min(1,\sigma_*+\rho(z))}\big)\big) \tag{1.12}$$

where $\sigma_* \doteq 1.457$ is defined at (12.9). Furthermore, (1.11) holds uniformly for $z < 1-\delta$ for any $\delta > 0$, and (1.12) holds uniformly for z in a compact subset of $(-\infty, 1)$.

Theorem 1.4 is proved in Section 12.1. It improves on bounds in [5, Section 2.7]. As a corollary of Theorem 1.4, we obtain a new proof of the following CLT. This CLT has been proved in [5, Theorem 1.7] via the recursion method (applied to moment generating functions) and in [2] via the martingale CLT. We prove Theorem 1.5 in Section 12.2.

Theorem 1.5.

$$\frac{D_n-\mu\log n}{\sqrt{\log n}} \xrightarrow{\mathrm{d}} \mathrm{Normal}(0,\sigma^2) \quad \textit{as } n\to\infty \tag{1.13}$$

where

$$\mu := 1/\zeta(2) = 6/\pi^2 \doteq 0.6079; \quad \sigma^2 := 2\zeta(3)/\zeta(2)^3 \doteq 0.5401. \tag{1.14}$$

Although not discussed in this article, there is a parallel CLT for L_n, which has been proved in several quite different ways. In [5, Theorem 1.7] using the recursion method. In [18] using the general *contraction method* [20]. In [16] using known results [11] in the theory of *regenerative composition structures*.

Another corollary of Theorem 1.4 is the following large deviation result, proved in Section 12.3.

Theorem 1.6. *As $n\to\infty$, we have:*

$$\mathbb{P}(D_n < x\log n) = n^{-\Lambda^*(x)+o(1)}, \qquad \textit{if} \quad 0 < x \le x_0, \tag{1.15}$$

$$\mathbb{P}(D_n > x\log n) = n^{-\Lambda^*(x)+o(1)}, \qquad \textit{if} \quad x_0 \le x < x_1, \tag{1.16}$$

$$\mathbb{P}(D_n > x\log n) \le n^{-\Lambda^*(x)+o(1)}, \qquad \textit{if} \quad x \ge x_1, \tag{1.17}$$

where $x_0 = 1/\zeta(2)$, $x_1 = 1/(\zeta(2)-1)$ and Λ^* is defined at (12.29).

Theorem 1.6 improves estimates for the upper tail in [5, Theorem 1.4] and [2].

Remark 1.7. Note that for L_n, we give a result only for the mean. We could not prove other results, for example for the variance, by the methods of this article (see [5] for results by the recurrence method). The difficulty is that for higher moments of L_n we have been unable to find a representation of the type in Proposition 4.1 and (11.1).

Applications beyond leaf heights. The results above concern only the two leaf heights D_n and L_n, but the methodology can also be applied to other aspects of the random tree model. We give two examples.

First consider $\Lambda_n :=$ (total) length of CTCS(n). Part of Theorem 9.1 is (see (9.9))

Theorem 1.8.

$$\mathbb{E}\left[\Lambda_n\right] = \frac{6}{\pi^2}n + O\bigl(n^{-|s_1|}\bigr). \tag{1.18}$$

A final application concerns the probability[3] $a(n,j)$ that the harmonic descent chain started at n ever visits state j. The following result was proved as Theorem 3.1 and (6.17) of [3], to give a first illustration of our methodology in that article.

Theorem 1.9. *For every fixed $j \geq 2$,*

$$a(n,j) = \frac{6}{\pi^2}\frac{h_{j-1}}{j-1} + O\bigl(n^{-1-|s_1|}\bigr). \tag{1.19}$$

In particular, as $n \to \infty$,

$$a(n,j) \to a(j) := \frac{6}{\pi^2}\frac{h_{j-1}}{j-1}. \tag{1.20}$$

1.4. **Outline of paper.** Section 2 describes the exchangeable partitions representation [3] of the limit CTCS(∞). Section 3 recalls some basic analysis surrounding Mellin transforms. Because the finite trees are embedded within CTCS(∞), several important expectations for the finite tree can be expressed in terms of a specific measure Υ defined for CTCS(∞) (Sections 4 - 5). The measure Υ is defined by its Mellin transform; we cannot invert explicitly but do understand its behavior near zero (Section 6). The remainder of the paper uses these tools to prove the Theorems stated above. This involves classical, but rather intricate, complex analysis.

2. The exchangeable partitions representation

The relation between trees and nested families of partitions has been used at least since [13]. Its application in the context of our model is explained in detail in [3], from which the material below is taken. See also [12] for closely related results in a greater generality.

Here we will consider CTCS(n). We do find it convenient to adopt the biological term *clade* for the set of leaves in a subtree.

Fix a level (time) $t \geq 0$. For each n, the clades of CTCS(n) at time t define a partition $\Pi^{[n]}(t)$ of $[n] := \{1, \ldots, n\}$. Now apply a uniform random permutation of

[3] This relates to the number of clades of size j within DTCS(n) – see (5.2) – and also to the study of the *fringe distribution* in [3].

$[n]$, so now the partition $\Pi^{[n]}(t)$ is exchangeable. Then, regarding CTCS(n) as a tree on leaves $[n]$, there is a natural "delete leaf n from the tree, and prune" operation [3, Section 2.3] that yields a tree with leaves $[n-1]$. There is a key non-obvious *consistency property*

> [3, Theorem 2.3] *The operation "delete leaf n from* CTCS(n) *and prune" gives a tree distributed as* CTCS($n-1$)*.*

So there exists a *consistent growth process* (CTCS(n)$, n \geq 1$) in which the partitions $\Pi^{[n]}(t)$ are consistent and therefore define a partition $\Pi(t)$ of $\mathbb{N} := \{1, 2, \dots\}$ into clades at time t. Explicitly, i and j (with $i, j \in \mathbb{N}$) are in the same part if and only if the branchpoint separating the paths to leaves i and j has height $> t$, in CTCS(n) for any $n \geq \max(i, j)$. In the sequel we consider only this consistent and exchangeable version of CTCS(n).

Because each CTCS(n) has been made exchangeable, $\Pi(t)$ is an exchangeable random partition of $\mathbb{N}$, so we can exploit the theory of exchangeable partitions. Denote the clades at time t, that is the parts of $\Pi(t)$, by $\Pi(t)_1, \Pi(t)_2, \dots$, enumerated in order of the least elements. In particular, the clade of leaf 1 is $\Pi(t)_1$. The clades $\Pi(t)_\ell$ are thus subsets of $\mathbb{N}$, and the clades of CTCS(n) are the sets $\Pi(t)_\ell \cap [n]$ that are non-empty. Note that $(\Pi_t)_{t \geq 0}$ determines CTCS(n) for every n.

Write $|\cdot|$ for cardinality. Define, for $\ell, n \geq 1$,

$$K_{t,\ell}^{(n)} := \big|\Pi(t)_\ell \cap [n]\big|; \tag{2.1}$$

the sequence $K_{t,1}^{(n)}, K_{t,2}^{(n)}, \dots$ is thus the sequence of sizes of the clades in CTCS(n), extended by 0's to an infinite sequence. By Kingman's fundamental result [6, Theorem 2.1], the asymptotic proportionate clade sizes, that is the limits

$$P_{t,\ell} := \lim_{n \to \infty} \frac{K_{t,\ell}^{(n)}}{n}, \tag{2.2}$$

exist a.s. for every $\ell \geq 1$, and the random partition $\Pi(t)$ may be reconstructed (in distribution) from the limits $(P_{t,\ell})_\ell$ by Kingman's paintbox construction, stated as the following theorem.

Theorem 2.1. *Let $t \geq 0$.*

(i) *If $t > 0$, then a.s. each $P_{t,\ell} \in (0,1)$, and $\sum_\ell P_{t,\ell} = 1$.*
(ii) *Given a realization of $(P_{t,\ell})_\ell$, give each integer $i \in \mathbb{N}$ a random colour ℓ, with probability distribution $(P_{t,\ell})_\ell$, independently for different i. These colours define a random partition $\Pi'(t)$ of $\mathbb{N}$, which has the same distribution as $\Pi(t)$.*

We note also that, as an immediate consequence of the paintbox construction, the distribution of $P_{t,1}$ equals the distribution of a size-biased sample from $\{P_{t,\ell}\}_{\ell=1}^{\infty}$ [6, Corollary 2.4]. In other words, for any function $f \geq 0$ on $(0,1)$ and $t > 0$,

$$\mathbb{E}\left[f(P_{t,1})\right] = \mathbb{E}\left[\sum_{\ell=1}^{\infty} P_{t,\ell} f(P_{t,\ell})\right] = \sum_{\ell=1}^{\infty} \mathbb{E}\left[P_{t,\ell} f(P_{t,\ell})\right]. \tag{2.3}$$

We have also [3, Theorem 4.5]

$$\mathbb{E}\left[P_{t,1}^s\right] = e^{-t(\psi(s+1)-\psi(1))}, \qquad \Re s > -1 \tag{2.4}$$

which will be key to our subsequent Mellin analysis.

3. Analysis preliminaries

3.1. **Asymptotic expansions.** An asymptotic expansion (see e.g. [10, p. 724]) of a function $f(n)$ written

$$f(n) \sim \sum_{k\geq 1} \lambda_k \omega_k(n), \tag{3.1}$$

for some functions $\omega_k(n)$ and real or complex coefficients λ_k $(k \geq 1)$, means that $\omega_{k+1}(n) = o(\omega_k(n))$ as $n \to \infty$ for every $k \geq 1$, and that for any $N \geq 1$, the error

$$f(n) - \sum_1^N \lambda_k \omega_k(n) = O\bigl(|\omega_{N+1}(n)|\bigr). \tag{3.2}$$

In other words, the error when approximating with a partial sum is of the order of the largest (non-zero) omitted term. Note that the infinite sum $\sum_1^\infty \lambda_k \omega_k(n)$ does not have to converge (and typically does not); this is indicated by the symbol "$\sim$" instead of "$=$" in (3.1).

In some expansions below we have several sums or a double sum in the asymptotic expansion; this should be interpreted as above by rearranging the terms in decreasing order.

3.2. **The digamma function.** The digamma function ψ is defined by [19, 5.2(i)]

$$\psi(z) := \frac{\mathrm{d}}{\mathrm{d}z}(\log\Gamma(z)) = \Gamma'(z)/\Gamma(z). \tag{3.3}$$

This is a meromorphic function in the complex plane $\mathbb{C}$, with simple poles at $\{0,-1,-2,\dots\}$. We have [19, 5.4.12]

$$\psi(1) = -\gamma, \qquad \psi'(1) = \zeta(2) = \tfrac{\pi^2}{6}, \qquad \psi''(1) = -2\zeta(3) \tag{3.4}$$

where $\gamma \doteq 0.5772$ is Euler's gamma, and more generally for integer $n \geq 1$ [19, 5.4.14]

$$\psi(n) = h_{n-1} + \psi(1) = h_{n-1} - \gamma. \tag{3.5}$$

We will use the asymptotic expansion [19, 5.11.2] (obtained by logarithmic differentiation of Stirling's formula), where B_k denotes the Bernoulli numbers,

$$\psi(z) \sim \log z - \frac{1}{2}z^{-1} - \sum_{k=1}^\infty \frac{B_{2k}}{2k} z^{-2k} \tag{3.6}$$

as $|z| \to \infty$ in any fixed sector $|\arg(z)| \leq \pi - \delta < \pi$. Note that, since it is valid in sectors, (3.6) may be termwise differentiated to yield asymptotic expansions of $\psi'(z)$

and higher derivatives, see [19, 5.15.8–9]. We have by (3.5) and (3.6) the asymptotic expansion

$$h_{n-1} \sim \log n + \gamma - \frac{1}{2}n^{-1} - \sum_{k=1}^{\infty} \frac{B_{2k}}{2k} n^{-2k}, \tag{3.7}$$

and in particular the well-known expansion

$$h_{n-1} = \log n + \gamma - \frac{1}{2n} + O\bigl(n^{-2}\bigr). \tag{3.8}$$

As the Mellin formula at (4.8) will show, the roots of $\psi(s) = \psi(1)$ will play an important role in our proofs and results, so we first describe them.

Lemma 3.1. *The roots of the equation $\psi(s) = \psi(1)$ are all real and can be enumerated in decreasing order as $s_0 = 1 > s_1 > s_2 > \dots$, with $s_i \in (-i, -(i-1))$ for $i \geq 1$. Numerically, $s_1 \doteq -0.567$ and $s_2 \doteq -1.628$.*

More generally, for any $a \in \mathbb{R}$, the roots of $\psi(s) = a$ are all real and can be enumerated as $s_0(a) = 1 > s_1(a) > s_2(a) > \dots$, with $s_i(a) \in (-i, -(i-1))$ for $i \geq 1$.

We let s_i have this meaning throughout the paper. See [19, §5.4(iii)] for $s_i(a)$ in the case $a = 0$.

Proof. Recall that $\psi(s)$ is a meromorphic function of s, with poles at $0, -1, -2, \dots$. For any other complex s we have the standard formulas [19, 5.7.6 and 5.15.1]

$$\psi(s) = -\gamma + \sum_{k=0}^{\infty} \Bigl(\frac{1}{k+1} - \frac{1}{k+s}\Bigr), \tag{3.9}$$

$$\psi'(s) = \sum_{k=0}^{\infty} \frac{1}{(k+s)^2}. \tag{3.10}$$

If $\Im s > 0$, then $\Im(1/(k+s)) < 0$ for all k and thus (3.9) implies $\Im \psi(s) > 0$. Similarly, if $\Im s < 0$, then $\Im \psi(s) < 0$. Consequently, all roots of $\psi(t) = a \in \mathbb{R}$ are real.

For real s, (3.10) shows that $\psi'(s) > 0$. We can write $\mathbb{R} \setminus \{\text{the poles}\} = \bigcup_{i=0}^{\infty} I_i$ with $I_0 := (0, \infty)$ and $I_i := (-i, -(i-1))$ for $i \geq 1$; it then follows that $\psi(s)$ is strictly increasing in each interval I_i. Moreover, by (3.9) (or general principles), at the poles we have the limits $\psi(-i-0) = +\infty$ and $\psi(-i+0) = -\infty$ $(i \geq 0)$, and furthermore (by (3.6)) $\psi(s) \to \infty$ as $s \nearrow +\infty$, so ψ maps each interval I_i to $(-\infty, \infty)$. Consequently, $\psi(s) = a$ has exactly one root $s_i(a)$ in each I_i. (See also the graph of $\psi(s)$ in [19, Figure 5.3.3].) In the special case $a = \psi(1)$, obviously the positive root is $s_0 = 1$.

The numerical values are obtained by `Maple`. ∎

3.3. **Mellin transforms.** If f is a function defined on $\mathbb{R}_+ := (0, \infty)$, then its *Mellin transform* is defined by

$$\widetilde{f}(s) := \int_0^{\infty} f(x) x^{s-1} \, \mathrm{d}x \tag{3.11}$$

for all complex s such that the integral converges absolutely. It is well known that the domain of all such s is a strip $\mathcal{D} = \{s : \Re s \in J_f\}$ for some interval $J_f \subseteq \mathbb{R}$ (possibly empty or degenerate), and that $\widetilde{f}(s)$ is analytic in the interior $\mathcal{D}^\circ$ of $\mathcal{D}$ (provided $\mathcal{D}^\circ$ is non-empty, i.e., J_f is neither empty nor degenerate).

In analogy to (3.11), if μ is a (possibly complex) measure on $\mathbb{R}_+$, then its *Mellin transform* is defined by

$$\widetilde{\mu}(s) := \int_0^\infty x^{s-1}\,\mathrm{d}\mu(x) \tag{3.12}$$

for all complex s such that the integral converges absolutely. Again, the domain of all such s is a strip $\mathcal{D} = \{s : \Re s \in J_\mu\}$ for some interval $J_\mu \subseteq \mathbb{R}$ (possibly empty or degenerate), and $\widetilde{\mu}(s)$ is analytic in the interior $\mathcal{D}^\circ$ of $\mathcal{D}$. Note that the Mellin transform (3.12) for real s is just the moments of the measure μ, with a simple shift of the argument.

Remark 3.2. There are in the literature also other definitions of the Mellin transform of measures; see for example the Mellin-Stieltjes transform in [8, Appendix D]. We use the definition above, which is convenient when we want to identify an absolutely continuous measure with its density function.

4. The central idea

Recall the limit $P_{t,1} := \lim_{n\to\infty} \frac{K_{t,1}^{(n)}}{n}$ from (2.2). The central idea in this article is to define an infinite measure Υ on $(0,1)$ by

$$\Upsilon := \int_0^\infty \mathcal{L}(P_{t,1})\,\mathrm{d}t. \tag{4.1}$$

This means that for any (measurable) function $f \geq 0$ on $(0,1)$,

$$\int_0^1 f(x)\,\mathrm{d}\Upsilon(x) = \int_0^\infty \int_0^1 f(x)\,\mathrm{d}\mathcal{L}(P_{t,1})(x)\,\mathrm{d}t = \int_0^\infty \mathbb{E}\left[f(P_{t,1})\right]\mathrm{d}t. \tag{4.2}$$

This extends by linearity to every complex-valued f such that $\int_0^\infty \mathbb{E}\left[|f(P_{t,1})|\right]\mathrm{d}t < \infty$. The identity (4.2) enables us to express several important expectations in terms of the measure Υ: these are listed in Proposition 4.1 below. As well as $\mathbb{E}[D_n]$ and $\mathbb{E}[L_n]$, these involve the total length Λ_n of CTCS(n), and the "occupation probability", that is

(4.3)
$a(n,i) :=$ probability that the harmonic descent chain started at state n is ever in state i.

So $a(n,n) = a(n,1) = 1$.

To illustrate the methodology we derive the first of these identities below (the others will be proved in Section 5).

Proposition 4.1. *For any $n \geq 1$:*

$$\mathbb{E}[D_n] = \int_0^1 \left(1 - (1-x)^{n-1}\right)\mathrm{d}\Upsilon(x). \tag{4.4}$$

$$a(n,j) = h_{j-1}\binom{n-1}{j-1}\int_0^1 x^{j-1}(1-x)^{n-j}\,\mathrm{d}\Upsilon(x), \quad 2 \le j \le n. \tag{4.5}$$

$$\mathbb{E}\left[L_n\right] = \sum_{j=2}^{n} a(n,j) = \int_0^1 \sum_{j=2}^{n} h_{j-1}\binom{n-1}{j-1}x^{j-1}(1-x)^{n-j}\,\mathrm{d}\Upsilon(x). \tag{4.6}$$

$$\mathbb{E}\left[\Lambda_n\right] = \int_0^1 \frac{1}{x}\Big(1-(1-x)^n - nx(1-x)^{n-1}\Big)\,\mathrm{d}\Upsilon(x). \tag{4.7}$$

But what is the measure Υ, explicitly? We first note that (4.2) and (2.4) tell us the Mellin transform of the measure Υ:

$$\widetilde{\Upsilon}(s) := \int_0^1 x^{s-1}\,\mathrm{d}\Upsilon(x) = \int_0^\infty \mathbb{E}\left[P_{t,1}^{s-1}\right]\mathrm{d}t = \frac{1}{\psi(s)-\psi(1)}, \qquad \Re s > 1. \tag{4.8}$$

The measure Υ is determined by its Mellin transform. We do not know how to invert the transform (4.8) to obtain a useful explicit formula for Υ, but what is relevant to our asymptotics is mainly the behavior of Υ near 0, which will be given in the "inversion estimate", Lemma 6.1.

4.1. **An illustrative identity.** To prove (4.4), consider CTCS(n) for an integer $n \ge 1$, and recall that the clades at time t are precisely the blocks of the partition $\Pi(t) \cap [n]$. Fix a non-empty subset $J \subseteq [n]$ of size $j = |J| \ge 1$. If we use the paintbox construction in Theorem 2.1 to reconstruct a copy $\Pi'(t)$ of $\Pi(t)$, we see that

> conditioned on $(P_{t,\ell})_{\ell=1}^\infty$, the probability that J is a block in $\Pi'(t)\cap[n]$ equals $\sum_\ell P_{t,\ell}^j(1-P_{t,\ell})^{n-j}$.

We take the expectation; $\Pi'(t)$ has the same distribution as $\Pi(t)$, so the distinction between them then disappears and we obtain that for any fixed non-empty set $J \subseteq [n]$ of size $|J| = j$ and $t > 0$,

$$\mathbb{P}\big(J \text{ is a clade at time } t\big) = \mathbb{E}\Big[\sum_\ell P_{t,\ell}^j(1-P_{t,\ell})^{n-j}\Big]. \tag{4.9}$$

By (2.3), we can rewrite (4.9) as

$$\mathbb{P}\big(J \text{ is a clade at time } t\big) = \mathbb{E}\left[P_{t,1}^{j-1}(1-P_{t,1})^{n-j}\right]. \tag{4.10}$$

In particular, since $D_n \le t$ exactly when $\{1\}$ is a clade at time t, (4.10) with $J = \{1\}$ and thus $j = 1$ yields

$$\mathbb{P}(D_n > t) = 1 - \mathbb{P}(D_n \le t) = 1 - \mathbb{E}\left[(1-P_{t,1})^{n-1}\right] = \mathbb{E}\left[1-(1-P_{t,1})^{n-1}\right]. \tag{4.11}$$

Consequently,

$$\mathbb{E}\left[D_n\right] = \int_0^\infty \mathbb{P}(D_n > t)\,\mathrm{d}t = \int_0^\infty \mathbb{E}\left[1-(1-P_{t,1})^{n-1}\right]\mathrm{d}t, \tag{4.12}$$

which by (4.2) yields our desired formula (4.4).

As with the other identities in Proposition 4.1, it is then straightforward (but sometimes intricate) classical analysis to combine (4.4) with the inversion estimate (Lemma 6.1) to obtain an asymptotic expansion of $\mathbb{E}[D_n]$: see Sections 7, 8, and 9.

Remark 4.2. We give in Section 10.1 a different argument, *not* using Lemma 6.1, where we obtain the same asymptotic expansion by directly combining (4.4) and the Mellin transform (4.8), using a version of Parseval's formula. This method is less intutive, but leads to exact formulas involving line integrals in the complex plane, which through residue calculus yield another proof of the asymptotic expansion.

Remark 4.3. We remark that (4.4) combined with (4.8) also yields an exact formula for $\mathbb{E}[D_n]$. By the binomial theorem we obtain, using (3.5),

$$\begin{aligned}\mathbb{E}[D_n] &= \int_0^1 \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n-1}{j} x^j \,\mathrm{d}\Upsilon(x) \\ &= \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n-1}{j} \frac{1}{\psi(j+1)-\psi(1)} = \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n-1}{j} \frac{1}{h_j}.\end{aligned} \tag{4.13}$$

However, since this is an alternating sum, it does not seem easy to derive asymptotics from it, and we will not use it.

5. Other identities involving Υ

5.1. **The occupation probability.** As noted in Section 1.2, if we follow the path from the root to leaf 1 (or equivalently to a uniformly random leaf) in CTCS(n) or DTCS(n), then the sequence of clade sizes follows a Markov chain, which we call the *harmonic descent* (HD) chain, see [2] and [4]. In particular, we are interested in the "occupation probability", that is

(5.1) $a(n,i) :=$ probability that the HD chain started at state n is ever in state i.

In other words, $a(n,i)$ is the probability that a given (or random) leaf in CTCS(n) or DTCS(n) belongs to some subtree with exactly j leaves. Writing $N_n(j)$ for the number of such subtrees, we clearly have

$$\mathbb{E}[N_n(j)] = na(n,j)/j. \tag{5.2}$$

Similar to the argument above that proved (4.4), one can prove [3, (6.7), (6.6), and (6.10)]the identity (4.5), which we repeat as

$$a(n,j) = h_{j-1}\binom{n-1}{j-1}\int_0^1 x^{j-1}(1-x)^{n-j}\,\mathrm{d}\Upsilon(x),\ 2\le j\le n. \tag{5.3}$$

The methodology via Lemma 6.1 that we use later was already applied in [3] to establish the asymptotics of $a(n,j)$ stated in Theorem 1.9, so we do not repeat it here.

Remark 5.1. Similarly to (4.13), we obtain from (5.3), (4.8), and (3.5) the exact formula, using a binomial expansion and manipulation of binomial coefficients,

$$\begin{aligned} a(n,j) &= h_{j-1}\binom{n-1}{j-1}\int_0^1 \sum_{k=0}^{n-j}\binom{n-j}{k} x^{j-1}(-x)^k \,\mathrm{d}\Upsilon(x) \\ &= h_{j-1}\sum_{k=0}^{n-j}\binom{n-1}{j-1}\binom{n-j}{k}\frac{(-1)^k}{\psi(j+k)-\psi(1)} \\ &= \sum_{k=0}^{n-j}(-1)^k\binom{n-1}{j-1,k,n-j-k}\frac{h_{j-1}}{h_{j+k-1}}. \end{aligned} \tag{5.4}$$

5.2. **The expected hop-height $\mathbb{E}[L_n]$.** In CTCS(n), the expected lifetime of a clade of size j is $1/h_{j-1}$, and hence (5.3) implies

$$\mathbb{E}[D_n] = \sum_{j=2}^{n}\frac{1}{h_{j-1}}a(n,j) = \int_0^1\sum_{j=2}^{n}\binom{n-1}{j-1}x^{j-1}(1-x)^{n-j}\,\mathrm{d}\Upsilon(x), \tag{5.5}$$

which by summing the binomial series yields another proof of (4.4). Similarly, for DTCS(n), the height L_n equals the number of clades of sizes ≥ 2 that leaf 1 ever belongs to, and thus we have

$$\mathbb{E}[L_n] = \sum_{j=2}^{n}a(n,j) = \int_0^1\sum_{j=2}^{n}h_{j-1}\binom{n-1}{j-1}x^{j-1}(1-x)^{n-j}\,\mathrm{d}\Upsilon(x), \tag{5.6}$$

We will proceed to the asymptotic expansion of $\mathbb{E}[L_n]$ in Section 8.

5.3. **The expected length $\mathbb{E}[\Lambda_n]$.** The number of edges of CTCS(n) equals $n-1$. Identifying *length* of an edge with *duration of time*, one can consider the length Λ_n of CTCS(n), that is the sum of all edge-lengths. In other words, Λ_n is the sum of the lifetimes of all internal nodes. Since there are $N_n(j)$ nodes with j descendants and each of them has an expected lifetime of $1/h_{j-1}$, with the lifetimes independent of $N_n(j)$, we obtain, using (5.2) and (5.3),

$$\begin{aligned} \mathbb{E}[\Lambda_n] &= \sum_{j=2}^{n}\frac{1}{h_{j-1}}\mathbb{E}[N_n(j)] = \sum_{j=2}^{n}\binom{n}{j}\int_0^1 x^{j-1}(1-x)^{n-j}\,\mathrm{d}\Upsilon(x) \\ &= \int_0^1\frac{1}{x}\Big(1-(1-x)^n-nx(1-x)^{n-1}\Big)\,\mathrm{d}\Upsilon(x). \end{aligned} \tag{5.7}$$

We will proceed to the asymptotic expansion of $\mathbb{E}[\Lambda_n]$ in Section 9.

6. The measure Υ

The following *inversion estimate* will allow us to pass from the identities in Proposition 4.1 to sharp $n\to\infty$ asymptotics for expectations. Part of this lemma is also given in [3]; for completeness, we give the entire proof.

Lemma 6.1. *Let Υ be the infinite measure on $(0,1)$ having the Mellin transform (4.8). Then Υ is absolutely continuous, with a continuous density $\upsilon(x)$ on $(0,1)$. Furthermore:*

(i) *The density $\upsilon(x)$ satisfies*

$$\upsilon(x) = \frac{6}{\pi^2 x} + O\bigl(x^{-s_1} + x^{-s_1}|\log x|^{-1}\bigr), \tag{6.1}$$

uniformly for $x \in (0,1)$, where $s_1 \doteq -0.567$ is the largest negative root of $\psi(s) = \psi(1)$. In other words, for $x \in (0,1)$,

$$\upsilon(x) = \frac{6}{\pi^2} x^{-1} + r(x) \tag{6.2}$$

where

$$r(x) = O\bigl(x^{-s_1} + x^{-s_1}|\log x|^{-1}\bigr) \tag{6.3}$$

and thus, in particular, for $x \in (0, \frac{1}{2})$ say,

$$r(x) = O\bigl(x^{|s_1|}\bigr). \tag{6.4}$$

Furthermore, $\Upsilon(\delta, 1) < \infty$ for every $\delta > 0$ and

$$\int_0^1 |r(x)| \,\mathrm{d}x < \infty. \tag{6.5}$$

(ii) *More generally, with $0 > s_1 > s_2 > \dots$ denoting the negative roots of $\psi(s) = \psi(1)$, for any $N \geq 0$,*

$$\upsilon(x) = \frac{6}{\pi^2} x^{-1} + \sum_{i=1}^{N} \frac{1}{\psi'(s_i)} x^{|s_i|} + r_N(x), \qquad 0 < x < 1, \tag{6.6}$$

where

$$r_N(x) = O\bigl(x^{|s_{N+1}|}(1 + |\log x|^{-1})\bigr) \tag{6.7}$$

and

$$\int_0^1 |r_N(x)| \,\mathrm{d}x < \infty. \tag{6.8}$$

By (6.4)–(6.5), the Mellin transform $\widetilde{r}(s)$ exists for $\Re s > s_1$, and we obtain by (6.2) and (4.8)

$$\widetilde{r}(s) = \widetilde{\Upsilon}(s) - \frac{6}{\pi^2}\int_0^1 x^{s-2}\,\mathrm{d}x = \frac{1}{\psi(s) - \psi(1)} - \frac{6}{\pi^2}\cdot\frac{1}{s-1}, \tag{6.9}$$

first for $\Re s > 1$ and then by analytic continuation for $\Re s > s_1$; note that the right-hand side of (6.9) has a removable singularity at $s = 1$ since the residues of the two terms cancel.

Proof of Lemma 6.1. We begin by noting, as in the proof of Lemma 3.1, that the Mellin transform $1/\bigl(\psi(s) - \psi(1)\bigr)$ in (4.8) extends to a meromorphic function in the entire complex plane, whose poles are the roots of $\psi(s) = \psi(1)$. As shown in Lemma 3.1, besides the obvious pole $s_0 = 1$, the other poles are real and negative, and thus can be ordered $0 > s_1 > s_2 > \dots$. In particular, there are no other poles than 1 in the half-plane $\Re s > s_1$, with $s_1 \doteq -0.567$. The residue at the pole $s_0 = 1$ is, using (3.4),

$$\operatorname{Res}_{s=1} \frac{1}{\psi(s)-\psi(1)} = \frac{1}{\psi'(1)} = \frac{6}{\pi^2}. \tag{6.10}$$

We cannot immediately use standard results on Mellin inversion (as in [9, Theorem 2(i)]) because the Mellin transform in (4.8) decreases too slowly as $\Im s \to \pm\infty$ to be integrable on a vertical line $\Re s = c$.[4] In fact, (3.6) implies that

$$\psi(s) = \log s + o(1) = \log|s| + O(1) = \log|\Im s| + O(1) \tag{6.11}$$

as $\Im s \to \infty$ with s in, for example, any half-plane $\Re s \geq c$.

We overcome this problem by differentiating the Mellin transform, but we first subtract the leading term corresponding to the pole at 1. Since Υ is an infinite measure, we first replace it by ν defined by $\mathrm{d}\nu(x) = x\,\mathrm{d}\Upsilon(x)$; note that ν is also a measure on $(0,1)$, and taking $s = 2$ in (4.8) shows that ν is a finite measure.

Next, define ν_0 as the measure $(6/\pi^2)\,\mathrm{d}x$ on $(0,1)$, and let ν_Δ be the (finite) signed measure $\nu - \nu_0$. Then ν_Δ has the Mellin transform, by (4.8),

$$\begin{aligned} \widetilde{\nu_\Delta}(s) &:= \int_0^1 x^{s-1}\,\mathrm{d}\nu_\Delta(x) = \int_0^1 x^s\,\mathrm{d}\Upsilon(x) - \frac{6}{\pi^2}\int_0^1 x^{s-1}\,\mathrm{d}x \\ &= \frac{1}{\psi(s+1)-\psi(1)} - \frac{6}{\pi^2 s}, \qquad \Re s > 0. \end{aligned} \tag{6.12}$$

We may here differentiate under the integral sign, which gives

$$\widetilde{\nu_\Delta}'(s) := \int_0^1 (\log x) x^{s-1}\,\mathrm{d}\nu_\Delta(x) \tag{6.13}$$

$$= -\frac{\psi'(s+1)}{(\psi(s+1)-\psi(1))^2} + \frac{6}{\pi^2 s^2}, \qquad \Re s > 0. \tag{6.14}$$

The Mellin transform $\widetilde{\nu_\Delta}(s)$ in (6.12) extends to a meromorphic function in $\mathbb{C}$ with (simple) poles $(s_i - 1)_1^\infty$; note that there is no pole at $s_0 - 1 = 0$, since the residues there of the two terms in (6.12) cancel by (6.10). Furthermore, the formula (6.14) for $\widetilde{\nu_\Delta}'(s)$ then holds for all s (although the integral in (6.13) diverges unless $\Re s > 0$).

For any real c we have, on the vertical line $\Re s = c$, as $\Im s \to \pm\infty$, that $\psi(s) \sim \log|s|$ by (6.11), and also, by differentiation of (6.11) (see [19, 5.15.8]) that $\psi'(s) \sim s^{-1}$. It follows from (6.14) that

$$\widetilde{\nu_\Delta}'(s) = O(|s|^{-1}\log^{-2}|s|) \tag{6.15}$$

on the line $\Re s = c$, for $|\Im s| \geq 2$ say, and thus $\widetilde{\nu_\Delta}'$ is integrable on this line unless c is one of the poles $s_i - 1$. In particular, taking $c = 1$ and thus $s = 1 + u\mathrm{i}$ $(u \in \mathbb{R})$, we see that the function

$$\widetilde{\nu_\Delta}'(1+\mathrm{i}u) = \int_0^1 x^{\mathrm{i}u}\log(x)\,\mathrm{d}\nu_\Delta(x) \tag{6.16}$$

[4] And we cannot use [9, Theorem 2(ii)] since we do not know that Υ has a density that is locally of bounded variation.

is integrable. The change of variables $x = e^{-y}$ shows that the function (6.16) is the Fourier transform of the signed measure on $\mathbb{R}_+$ that corresponds to $\log(x)\,\mathrm{d}\nu_\Delta(x)$. This measure on $\mathbb{R}_+$ is thus a finite signed measure with integrable Fourier transform, which implies that it is absolutely continuous with a continuous density. Reversing the change of variables, we thus see that the signed measure $\log(x)\,\mathrm{d}\nu_\Delta(x)$ is absolutely continuous with a continuous density on $(0,1)$. Moreover, denoting this density by $h(x)$, we obtain the standard inversion formula for the Mellin transform [9, Theorem 2(i)], [19, 1.14.35]:

$$h(x) = \frac{1}{2\pi \mathrm{i}} \int_{c-\infty\mathrm{i}}^{c+\infty\mathrm{i}} x^{-s} \widetilde{\nu_\Delta}'(s)\,\mathrm{d}s, \qquad x > 0, \tag{6.17}$$

with $c = 1$. Furthermore, the integrand in (6.17) is analytic in the half-plane $\Re s > s_1 - 1$, and the estimate (6.15) above is uniform for $\Re s$ in any compact interval and $|\Im s| \geq 2$. Consequently, we may shift the line of integration in (6.17) to any $c > s_1 - 1$. Taking absolute values in (6.17), and recalling that $\widetilde{\nu_\Delta}'(s)$ is integrable on the line, then yields

$$h(x) = O\bigl(x^{-c}\bigr) \tag{6.18}$$

for any $c > s_1 - 1$.

Reversing the transformations above, we see that ν_Δ has the density $(\log x)^{-1} h(x)$, and thus ν has the density $(\log x)^{-1} h(x) + 6/\pi^2$, and, finally, that Υ has the density

$$\upsilon(x) := \frac{\mathrm{d}\Upsilon}{\mathrm{d}x} = \frac{1}{x}\frac{\mathrm{d}\nu}{\mathrm{d}x} = \frac{6}{\pi^2 x} + \frac{1}{x \log x} h(x), \qquad 0 < x < 1. \tag{6.19}$$

Furthermore, (6.19) and (6.18) have the form of the claimed estimate (6.1), although with the weaker error term $O(x^{-s_1-\varepsilon}|\log x|^{-1})$ for any $\varepsilon > 0$.

To obtain the claimed error term, we note that the residue of $\widetilde{\nu_\Delta}(s)$ at $s_1 - 1$ is $a_1 := 1/\psi'(s_1)$. Let ν_1 be the measure $a_1 x^{1-s_1}\,\mathrm{d}x$ on $(0,1)$; then ν_1 has Mellin transform

$$\widetilde{\nu_1}(s) = a_1 \int_0^1 x^{s-1} x^{1-s_1}\,\mathrm{d}x = \frac{a_1}{s + 1 - s_1}, \qquad \Re s > s_1 - 1. \tag{6.20}$$

It follows from (6.12) and (6.20) that the signed measure $\nu - \nu_0 - \nu_1 = \nu_\Delta - \nu_1$ has the Mellin transform

$$\frac{1}{\psi(s+1) - \psi(1)} - \frac{6}{\pi^2 s} - \frac{a_1}{s + 1 - s_1}, \tag{6.21}$$

which is an analytic function in the half plane $\Re s > s_2 - 1$. Hence, the same argument as above yields the estimate

$$\upsilon(x) = \frac{6}{\pi^2 x} + \frac{1}{\psi'(s_1)} x^{-s_1} + O\bigl(x^{-s_2-\varepsilon}|\log x|^{-1}\bigr), \qquad x \downarrow 0, \tag{6.22}$$

for any $\varepsilon > 0$, which in particular yields (6.1), and thus (6.2)–(6.3).

We may continue the argument above further and subtract similar terms for any number of poles; this leads to the estimate, for any $N \geq 1$,

$$\upsilon(x) = \frac{6}{\pi^2 x} + \sum_{i=1}^{N} \frac{1}{\psi'(s_i)} x^{-s_i} + O\big(x^{-s_{N+1}}(1 + |\log x|^{-1})\big), \qquad 0 < x < 1, \tag{6.23}$$

and thus (6.6)–(6.7).

Finally, $\Upsilon(\delta, 1) < \infty$ for $\delta > 0$ follows directly from (4.8) with $s = 2$, say. Hence, (6.2) implies $\int_{1/2}^{1} |r(x)| \, \mathrm{d}x < \infty$, while $\int_0^{1/2} |r(x)| \, \mathrm{d}x < \infty$ follows from (6.4); thus (6.5) holds. The same argument yields (6.8). □

Remark 6.2. Although the function $h(x)$ is continuous also at $x = 1$ (and thus $h(1) = 0$), the density $\upsilon(x)$ diverges as $x \nearrow 1$ because of the factor $\log x$ in the denominator in (6.19); in fact, it can be shown by similar arguments that

$$\upsilon(1 - y) \sim \frac{1}{y \, |\log y|^2}, \qquad y \searrow 0. \tag{6.24}$$

Hence, $r(x)$ and $r_N(x)$ are unbounded on $(0, 1)$ and the error terms in (6.1), (6.3), and (6.7) cannot be simplified as in (6.4) on the entire interval $(0, 1)$.

7. The expected leaf height $\mathbb{E}[D_n]$

Armed with Lemma 6.1, we may now, as said in Section 4, obtain an asymptotic expansion of $\mathbb{E}[D_n]$ from (4.4), which extends the first terms given in [5, Theorem 1.1]. We begin by finding the leading terms, using (6.1)–(6.3) only: that gives Proposition 7.2. We then extend this to a full asymptotic expansion.

7.1. **Leading terms.**

We denote the integrand in (4.4) by

$$f_n(x) := \begin{cases} 1 - (1 - x)^{n-1}, & 0 < x < 1, \\ 0, & x \geq 1. \end{cases} \tag{7.1}$$

In (4.4), we also recall that $\mathrm{d}\Upsilon(x) = \upsilon(x) \, \mathrm{d}x$ and substitute $\upsilon(x)$ using (6.2); this yields two terms:

$$\mathbb{E}[D_n] = \int_0^1 f_n(x) \frac{6}{\pi^2 x} \, \mathrm{d}x + \int_0^1 f_n(x) r(x) \, \mathrm{d}x. \tag{7.2}$$

For the first (main) term we use the following lemma (stated in a general form for later use), which gives the Mellin transform of f_n.

Lemma 7.1. *Fix $n \geq 1$. The Mellin transform*

$$\widetilde{f}_n(s) := \int_0^1 x^{s-1}\big(1 - (1 - x)^{n-1}\big) \, \mathrm{d}x, \qquad \Re s > -1, \tag{7.3}$$

is analytic in the half-plane $\Re s > -1$ and is given explicitly by

$$\widetilde{f}_n(s) = \frac{1}{s} - \frac{\Gamma(s)\Gamma(n)}{\Gamma(n + s)} = \frac{1}{s}\Big(1 - \frac{\Gamma(s + 1)\Gamma(n)}{\Gamma(n + s)}\Big), \qquad s \neq 0, \tag{7.4}$$

$$\widetilde{f}_n(0) = h_{n-1}. \tag{7.5}$$

Proof. It is elementary that the integral in (7.3) converges and defines an analytic (in fact, rational) function for $\Re s > -1$. For $\Re s > 0$, we may write $\widetilde{f}_n(s) = \int_0^1 x^{s-1}\,\mathrm{d}x - \int_0^1 x^{s-1}(1-x)^{n-1}\,\mathrm{d}x$, and (7.4) follows by evaluating the beta integral. Hence, (7.4) follows by analytic continuation.

For $s = 0$, the right-hand side of (7.4) has a removable singularity, with the value, by the definition of derivative (or L'Hôpital's rule), $\Gamma'(s) = \psi(s)\Gamma(s)$, and (3.5),

$$\widetilde{f}_n(0) = -\frac{\mathrm{d}}{\mathrm{d}s}\frac{\Gamma(s+1)\Gamma(n)}{\Gamma(n+s)}\Big|_{s=0} = -\psi(1) + \psi(n) = h_{n-1}, \tag{7.6}$$

which shows (7.5). □

By (4.4), (6.2), and Lemma 7.1, the main term of $\mathbb{E}\,[D_n]$ in (7.2) is

$$\int_0^1 \bigl[1-(1-x)^{n-1}\bigr]\frac{6}{\pi^2 x}\,\mathrm{d}x = \frac{6}{\pi^2}\widetilde{f}_n(0) = \frac{6}{\pi^2}h_{n-1}. \tag{7.7}$$

For the remainder term in (7.2), we use (6.4) and (6.5). Consequently,

$$\begin{aligned}
&\int_0^1 \bigl[1-(1-x)^{n-1}\bigr]r(x)\,\mathrm{d}x \\
&\qquad = \int_0^1 r(x)\,\mathrm{d}x - \int_0^{1/2}(1-x)^{n-1}r(x)\,\mathrm{d}x - \int_{1/2}^1 (1-x)^{n-1}r(x)\,\mathrm{d}x \\
&\qquad = \widetilde{r}(1) + \int_0^{1/2}(1-x)^{n-1}O(x^{|s_1|})\,\mathrm{d}x + \int_{1/2}^1 O\bigl(2^{-n}|r(x)|\bigr)\,\mathrm{d}x \\
&\qquad = \widetilde{r}(1) + O(n^{-1-|s_1|}) + O(2^{-n}) = \widetilde{r}(1) + O(n^{-1-|s_1|}),
\end{aligned} \tag{7.8}$$

where we estimate the penultimate integral using $(1-x)^{n-1} \le e^{-(n-1)x}$ (or by another beta integral). Combining (4.4), (6.2), (7.7) and (7.8) we obtain

$$\mathbb{E}\,[D_n] = \frac{6}{\pi^2}h_{n-1} + \widetilde{r}(1) + O(n^{-1-|s_1|}), \tag{7.9}$$

which by (3.8) yields the desired leading terms:

Proposition 7.2.

$$\mathbb{E}\,[D_n] = \frac{6}{\pi^2}\log n + c_0 + c_{-1}n^{-1} + O(n^{-1-|s_1|}) \tag{7.10}$$

for $c_0 := \widetilde{r}(1) + \frac{6}{\pi^2}\gamma$ *and* $c_{-1} = -\frac{3}{\pi^2}$.

In particular, this verifies the conjectured formula [5, (2.17)]. Moreover, we may compute the constants $\widetilde{r}(1)$ above by taking the limit as $s \to 1$ in (6.9). A simple calculation (using a Taylor expansion of $\psi(1+\varepsilon) - \psi(1)$) yields

$$\widetilde{r}(1) = -\frac{\psi''(1)}{2(\psi'(1))^2} = \frac{\zeta(3)}{\zeta(2)^2}, \tag{7.11}$$

using $\psi'(1) = \zeta(2)$ and $\psi''(1) = -2\zeta(3)$, see (3.4) and (3.10) or [19, 5.7.4]. Hence,

$$c_0 = \frac{\zeta(3)}{\zeta(2)^2} + \frac{\gamma}{\zeta(2)} \doteq 0.795155660439 \tag{7.12}$$

which agrees to 10 decimals with the numerical estimate in [5].

7.2. **A full asymptotic expansion.** The expansion (7.9) is easily extended to a full asymptotic expansion of $\mathbb{E}[D_n]$, stated in the introduction as Theorem 1.1.

Theorem 7.3. *We have the asymptotic expansion*

$$\mathbb{E}[D_n] \sim \frac{6}{\pi^2} h_{n-1} + \frac{\zeta(3)}{\zeta(2)^2} - \sum_{i=1}^{\infty} \frac{\Gamma(|s_i|+1)}{\psi'(s_i)} \frac{\Gamma(n)}{\Gamma(n+|s_i|+1)} \tag{7.13}$$

where $0 > s_1 > s_2 > \dots$ *are the negative roots of* $\psi(s) = \psi(1)$*. Alternatively, we have*

$$\mathbb{E}[D_n] \sim \frac{6}{\pi^2} \log n + \sum_{i=0}^{\infty} c_i n^{-i} + \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} c_{j,k}\, n^{-|s_j|-k} \tag{7.14}$$

for some coefficients c_i *and* $c_{j,k}$ *that can be found explicitly; in particular,* c_0 *is given by* (7.12) *and* $c_1 = -3/\pi^2$*.*

Proof. We use again (7.2), but we now note also that (6.2) and (6.6) yield

$$r(x) = \sum_{i=1}^{N} \frac{1}{\psi'(s_i)} x^{|s_i|} + r_N(x), \qquad 0 < x < 1. \tag{7.15}$$

Hence, similarly to (7.8) but now using also (6.7) and (6.8) and evaluating beta integrals,

$$\begin{aligned} &\int_0^1 \bigl[1-(1-x)^{n-1}\bigr] r(x)\,\mathrm{d}x \\ &= \int_0^1 r(x)\,\mathrm{d}x - \sum_{i=1}^{N} \frac{1}{\psi'(s_i)} \int_0^1 x^{|s_i|}(1-x)^{n-1}\,\mathrm{d}x - \int_0^1 (1-x)^{n-1} r_N(x)\,\mathrm{d}x \\ &= \widetilde{r}(1) - \sum_{i=1}^{N} \frac{1}{\psi'(s_i)} \frac{\Gamma(|s_i|+1)\Gamma(n)}{\Gamma(n+|s_i|+1)} + \int_0^{1/2} (1-x)^{n-1} O(x^{|s_{N+1}|})\,\mathrm{d}x \\ &\qquad\qquad + \int_{1/2}^1 O\bigl(2^{-n}|r_N(x)|\bigr)\,\mathrm{d}x \\ &= \widetilde{r}(1) - \sum_{i=1}^{N} \frac{\Gamma(|s_i|+1)}{\psi'(s_i)} \frac{\Gamma(n)}{\Gamma(n+|s_i|+1)} + O(n^{-1-|s_{N+1}|}). \end{aligned} \tag{7.16}$$

Using (7.16) instead of (7.8) in combination with (4.4), (6.2), (7.7), and (7.11) yields

$$\mathbb{E}[D_n] = \frac{6}{\pi^2} h_{n-1} + \frac{\zeta(3)}{\zeta(2)^2} - \sum_{i=1}^{N} \frac{\Gamma(|s_i|+1)}{\psi'(s_i)} \frac{\Gamma(n)}{\Gamma(n+|s_i|+1)} + O(n^{-1-|s_{N+1}|}). \tag{7.17}$$

Since N is arbitrary, this shows the asymptotic expansion (7.13).

Finally, (7.14) follows from (7.13). In fact, for every fixed b, we have the asymptotic expansion [19, 5.11.13]

$$\frac{\Gamma(n)}{\Gamma(n+b)} \sim \sum_{k=0}^{\infty} g_k(b) n^{-b-k}, \tag{7.18}$$

for some coefficients $g_k(b)$ that can be calculated explicitly [19, 5.11.15 and 17]. Substituting these expansions and (3.7) in (7.13) yields (7.14). $\square$

Example 7.4. Recall from Lemma 3.1 that $s_1 \doteq -0.567$ and $s_2 \doteq -1.628$. The first terms in (7.14) thus yield, with c_0 given by (7.12) and $c_2 = -1/(2\pi^2)$ by (3.7),

$$\mathbb{E}\left[D_n\right] = \frac{6}{\pi^2}\log n + c_0 - \frac{3}{\pi^2}n^{-1} - \frac{\Gamma(|s_1|+1)}{\psi'(s_1)}n^{-|s_i|-1} - \frac{1}{2\pi^2}n^{-2} + O\bigl(n^{-|s_1|-2}\bigr), \tag{7.19}$$

where $-|s_1|-1 \doteq -1.567$. The next terms are constants times $n^{-|s_1|-2}$ and $n^{-|s_2|-1}$, with exponents -2.567 and -2.628. Numerically, the coefficient of $n^{-|s_1|-1}$ is $-\Gamma(|s_1|+1)/\psi'(s_1) \doteq -0.0943$.

8. The expected hop-height $\mathbb{E}\left[L_n\right]$

We next find a similar asymptotic expansion for the expectation of the hop-height L_n in discrete time, extending the first terms given in [5, Theorem 1.2] by a different method.

Theorem 8.1. *We have the asymptotic expansion*

$$\mathbb{E}\left[L_n\right] \sim \frac{3}{\pi^2}h_{n-1}^2 + \frac{\zeta(3)}{\zeta(2)^2}h_{n-1} + \frac{\zeta(3)^2}{\zeta(2)^3} + \frac{1}{10} - \frac{3}{\pi^2}\psi'(n) + \sum_{i=1}^{\infty} \frac{\Gamma(|s_i|+1)}{(|s_i|+1)\psi'(s_i)} \frac{\Gamma(n)}{\Gamma(n+|s_i|+1)} \tag{8.1}$$

where $0 > s_1 > s_2 > \dots$ are the negative roots of $\psi(s) = \psi(1)$. Alternatively, we have

$$\mathbb{E}\left[L_n\right] \sim \frac{3}{\pi^2}\log^2 n + \Bigl(\frac{\zeta(3)}{\zeta(2)^2} + \frac{\gamma}{\zeta(2)}\Bigr)\log n + b_0 + \sum_{k=1}^{\infty} a_k n^{-k}\log n + \sum_{k=1}^{\infty} b_k n^{-k} + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} c_{j,k} n^{-|s_j|-k} \tag{8.2}$$

for some computable constants a_k, b_k, $c_{j,k}$; in particular,

$$b_0 = \frac{3\gamma^2}{\pi^2} + \frac{\zeta(3)}{\zeta(2)^2}\gamma + \frac{\zeta(3)^2}{\zeta(2)^3} + \frac{1}{10} \doteq 0.78234. \tag{8.3}$$

Remark 8.2. The coefficient for $\log n$ in (8.2) (found already in [5]) equals the constant term c_0 in the asymptotic expansion (7.14) of $\mathbb{E}\left[D_n\right]$.

Proof. We write (4.6) as

$$\mathbb{E}\,[L_n] = \int_0^1 H_n(x)\,\mathrm{d}\Upsilon(x), \tag{8.4}$$

where we (substituting $j = k+1$) define the function,

$$H_n(x) := \sum_{k=1}^{n-1} h_k \binom{n-1}{k} x^k (1-x)^{n-1-k}, \qquad 0 \le x \le 1. \tag{8.5}$$

To obtain a more tractable form of H_n for our analysis, we note that

$$h_k = \sum_{i=1}^{k} \frac{1}{i} = \sum_{i=1}^{k} \int_0^1 u^{i-1}\,\mathrm{d}u = \int_0^1 \frac{1-u^k}{1-u}\,\mathrm{d}u. \tag{8.6}$$

Hence, (8.5) yields (for convenience shifting the index to $n+1$)

$$\begin{aligned} H_{n+1}(x) &= \sum_{k=1}^{n} \int_0^1 \mathrm{d}u \frac{1-u^k}{1-u} \binom{n}{k} x^k (1-x)^{n-k} \\ &= \int_0^1 \frac{\mathrm{d}u}{1-u} \sum_{k=0}^{n} \binom{n}{k} \bigl(x^k - (ux)^k\bigr)(1-x)^{n-k} \\ &= \int_0^1 \frac{\mathrm{d}u}{1-u} \bigl(1 - (1-x+ux)^n\bigr) \\ &= \int_0^1 \bigl(1-(1-xw)^n\bigr)\frac{\mathrm{d}w}{w} \\ &= \int_0^x \bigl(1-(1-y)^n\bigr)\frac{\mathrm{d}y}{y}. \end{aligned} \tag{8.7}$$

Since (8.5) yields $H_n(1) = h_{n-1}$ (which also follows by (8.7) and (7.5)), (8.7) yields also

$$H_n(x) = H_n(1) - \int_x^1 \bigl(1-(1-y)^{n-1}\bigr)\frac{\mathrm{d}y}{y} = h_{n-1} - \int_x^1 \bigl(1-(1-y)^{n-1}\bigr)\frac{\mathrm{d}y}{y}. \tag{8.8}$$

By (8.5), $H_n(x)$ is a polynomial on $[0,1]$ with $H_n(0) = 0$, and thus the Mellin transform $\widetilde{H}_n(s)$ exists for $\Re s > -1$. When $\Re s > 0$, we have by (8.8) and (7.3)

$$\begin{aligned} \widetilde{H}_n(s) &= \int_0^1 x^{s-1} H_n(x)\,\mathrm{d}x \\ &= \frac{h_{n-1}}{s} - \iint_{0<x<y<1} x^{s-1}\bigl(1-(1-y)^{n-1}\bigr)y^{-1}\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{h_{n-1}}{s} - \frac{1}{s}\int_{0<y<1} y^{s-1}\bigl(1-(1-y)^{n-1}\bigr)\,\mathrm{d}y \\ &= \frac{1}{s}\bigl(h_{n-1} - \widetilde{f}_n(s)\bigr), \end{aligned} \tag{8.9}$$

and this extends to $\Re s > -1$ by analytic continuation. (The right-hand side is analytic in this domain by Lemma 7.1 and (7.5).) For $s = 0$, the right-hand side of (8.9) is interpreted as a limit in the standard way, giving

$$\widetilde{H}_n(0) = -\widetilde{f}_n'(0). \tag{8.10}$$

We now use (8.4), (8.7), and the expansion (6.6) for the density $f(x)$ of Υ. First, by (8.4) and (6.2),

$$\mathbb{E}\,[L_n] = \frac{6}{\pi^2}\int_0^1 H_n(x)\frac{1}{x}\,\mathrm{d}x + \int_0^1 H_n(x)r(x)\,\mathrm{d}x. \tag{8.11}$$

For the main term, we note that $\int_0^1 H_n(x)x^{-1}\,\mathrm{d}x = \widetilde{H}_n(0)$, which by (8.10) equals $-\widetilde{f}_n'(0)$. To compute this derivative, we use (7.4) and make a Taylor expansion of $\Gamma(s+1)\Gamma(n)/\Gamma(n+s)$ to obtain, recalling (3.3) and (3.4)–(3.5),

$$\begin{aligned}\int_0^1 H_n(x)\frac{1}{x}\,\mathrm{d}x &= \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\frac{\Gamma(s+1)\Gamma(n)}{\Gamma(n+s)}\Big|_{s=0} \\ &= \tfrac{1}{2}\big[\big(\psi(1)-\psi(n)\big)^2 + \psi'(1) - \psi'(n)\big] \\ &= \tfrac{1}{2}h_{n-1}^2 + \tfrac{\pi^2}{12} - \tfrac{1}{2}\psi'(n).\end{aligned} \tag{8.12}$$

For the final term in (8.11), we use (8.8) and obtain, recalling (7.11),

$$\begin{aligned}&\int_0^1 H_n(x)r(x)\,\mathrm{d}x \\ &= h_{n-1}\int_0^1 r(x)\,\mathrm{d}x - \iint_{0<x<y<1}\frac{1-(1-y)^{n-1}}{y}r(x)\,\mathrm{d}x\,\mathrm{d}y \\ &= \widetilde{r}(1)h_{n-1} - \iint_{0<x<y<1}\frac{1}{y}r(x)\,\mathrm{d}x\,\mathrm{d}y + \iint_{0<x<y<1}\frac{(1-y)^{n-1}}{y}r(x)\,\mathrm{d}x\,\mathrm{d}y \\ &= \widetilde{r}(1)h_{n-1} + \int_0^1(\log x)r(x)\,\mathrm{d}x + \int_0^1\frac{(1-y)^{n-1}}{y}\int_0^y r(x)\,\mathrm{d}x\,\mathrm{d}y.\end{aligned} \tag{8.13}$$

A differentiation under the integral sign in (3.11) shows that

$$\int_0^1(\log x)r(x)\,\mathrm{d}x = \widetilde{r}\,'(1). \tag{8.14}$$

For the final integral in (8.13) we use the expansion (7.15). Note that a term x^s $(s \ge 0)$ in $r(x)$ when substituted into this integral yields

$$\frac{1}{s+1}\int_0^1\frac{(1-y)^{n-1}}{y}y^{s+1}\,\mathrm{d}y = \frac{\Gamma(s+1)\Gamma(n)}{(s+1)\Gamma(n+s+1)}. \tag{8.15}$$

For the remainder term r_N, we split the integral into two parts as in (7.16), and use (6.7) for $y \in (0, \frac{1}{2})$ and (6.8) for $y \in (\frac{1}{2}, 1)$. Hence, (8.13) and (7.15) yield

$$\begin{aligned}(8.16) \qquad & \int_0^1 H_n(x) r(x)\,\mathrm{d}x \\ &= \widetilde{r}(1) h_{n-1} + \widetilde{r}\,'(1) + \sum_{i=1}^{N} \frac{1}{\psi'(s_i)} \frac{\Gamma(|s_i|+1)\Gamma(n)}{(|s_i|+1)\Gamma(n+|s_i|+1)} + O\bigl(n^{-|s_{N+1}|-1}\bigr).\end{aligned}$$

To find $\widetilde{r}\,'(1)$ we use (6.9) and find by a Taylor expansion of $\psi(s)$, using $\psi'(1) = \zeta(2) = \pi^2/6$, $\psi''(1) = -2\zeta(3)$ and $\psi'''(1) = 6\zeta(4) = \pi^4/15$ (see (3.4) and (3.10) or [19, 5.7.4]),

$$(8.17) \qquad \widetilde{r}\,'(1) = \frac{(\psi''(1))^2}{4(\psi'(1))^3} - \frac{\psi'''(1)}{6(\psi'(1))^2} = \frac{\zeta(3)^2}{\zeta(2)^3} - \frac{\zeta(4)}{\zeta(2)^2} = \frac{\zeta(3)^2}{\zeta(2)^3} - \frac{2}{5}.$$

We obtain (8.1) by (8.11), (8.12), (8.16), (7.11), and (8.17).

Finally, we obtain (8.2) from (8.1) by substituting (3.7), the corresponding asymptotic expansion of $\psi'(n)$ (obtained from (3.6) by termwise differentiation), and (7.18), and then rearranging the terms. $\square$

9. The length of CTCS(n)

Recall from (4.7) that the length Λ_n of CTCS(n) satisfies the identity

$$(9.1) \qquad \mathbb{E}\,[\Lambda_n] = \int_0^1 \frac{1}{x}\Bigl(1 - (1-x)^n - nx(1-x)^{n-1}\Bigr)\,\mathrm{d}\Upsilon(x).$$

We denote the integrand in the integral by

$$(9.2) \qquad \lambda_n(x) := \bigl(1 - (1-x)^n - nx(1-x)^{n-1}\bigr)/x.$$

Then its Mellin transform is, by beta integrals and simple algebra,

$$\begin{aligned}(9.3) \qquad \widetilde{\lambda}_n(s) &= \int_0^1 \Bigl(x^{s-2} - x^{s-2}(1-x)^n - nx^{s-1}(1-x)^{n-1}\Bigr)\,\mathrm{d}x \\ &= \frac{1}{s-1} - \frac{\Gamma(s-1)\Gamma(n+1)}{\Gamma(n+s)} - n\frac{\Gamma(s)\Gamma(n)}{\Gamma(n+s)} \\ &= \frac{1}{s-1}\Bigl(1 - \frac{\Gamma(s+1)\Gamma(n+1)}{\Gamma(n+s)}\Bigr),\end{aligned}$$

first assuming $\Re s > 1$ and then by analytic continuation for $\Re s > -1$ (the domain where $\widetilde{\lambda}_n(s)$ exist, for any $n \geq 2$); note that $s = 1$ is a removable singularity in (9.3) and *not* a pole. In particular, (9.3) yields

$$(9.4) \qquad \widetilde{\lambda}_n(0) = -(1-n) = n - 1.$$

We use Lemma 6.1 and obtain from (6.2), (9.1), and (9.2)

$$(9.5) \qquad \mathbb{E}\,[\Lambda_n] = \int_0^1 \lambda_n(x) \upsilon(x)\,\mathrm{d}x = \frac{6}{\pi^2}\int_0^1 \lambda_n(x) x^{-1}\,\mathrm{d}x + \int_0^1 \lambda_n(x) r(x)\,\mathrm{d}x.$$

The main term here is, using (9.4),

$$\frac{6}{\pi^2}\int_0^1 \lambda_n(x)x^{-1}\,\mathrm{d}x = \frac{6}{\pi^2}\widetilde{\lambda}_n(0) = \frac{6}{\pi^2}(n-1) \tag{9.6}$$

while the remainder term in (9.4) can be estimated as, by arguments as in (7.8),

$$\begin{aligned}\int_0^1 \lambda_n(x)r(x)\,\mathrm{d}x \\ &= \int_0^1 x^{-1}r(x)\,\mathrm{d}x - \int_0^1 (1-x)^n x^{-1}r(x)\,\mathrm{d}x - n\int_0^1 (1-x)^{n-1}r(x)\,\mathrm{d}x \\ &= \widetilde{r}(0) + O\big(n^{-|s_1|}\big).\end{aligned} \tag{9.7}$$

Furthermore, $\psi(s)$ has a pole at $s=0$, and thus (6.9) yields

$$\widetilde{r}(0) = 0 - \frac{6}{\pi^2}\cdot\frac{1}{-1} = \frac{6}{\pi^2}. \tag{9.8}$$

Consequently, by (9.5)–(9.8),

$$\mathbb{E}\,[\Lambda_n] = \frac{6}{\pi^2}(n-1) + \frac{6}{\pi^2} + O\big(n^{-|s_1|}\big) = \frac{6}{\pi^2}n + O\big(n^{-|s_1|}\big). \tag{9.9}$$

We may easily extend the argument above and obtain the following full asymptotic expansion.

Theorem 9.1.

$$\mathbb{E}\,[\Lambda_n] \sim \frac{6}{\pi^2}n - \sum_{i=1}^{\infty}\frac{\Gamma(|s_i|+2)}{|s_i|\psi'(s_i)}\,\frac{\Gamma(n+1)}{\Gamma(n+|s_i|+1)}. \tag{9.10}$$

Alternatively, for some coefficients $c_{i,k}$ that can be found explicitly,

$$\mathbb{E}\,[\Lambda_n] \sim \frac{6}{\pi^2}n + \sum_{j=1}^{\infty}\sum_{k=0}^{\infty} c_{j,k}\, n^{-|s_j|-k} \tag{9.11}$$

Note that the terms in the sum (9.10) have orders $n^{-|s_i|}$.

Proof. This is similar to previous proofs, so we omit some details.

We may use (7.15) in the last two integrals in (9.7); then calculations similar to (7.16) yield (9.10). Finally, (9.10) implies (9.11) by (7.18). □

10. Alternative proofs via a Parseval formula

We now show that one may skip the intermediate step of obtaining asymptotics for the measure Υ by using the following version of Parseval's formula (also called Plancherel's formula) for Mellin transforms: see Appendix A for a proof and references.

Recall that an integral $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$ exists *conditionally* if f is locally integrable and the symmetric limit $\lim_{A\to\infty}\int_{-A}^{A} f(x)\,\mathrm{d}x$ exists (and is finite); this limit is then defined to be the integral $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$. We use the same terminology for line integrals $\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}$.

Lemma 10.1. *Suppose that f is a locally integrable function and μ a measure on $\mathbb{R}_+$, and that $\sigma \in \mathbb{R}$ is such that $\int_0^\infty x^{\sigma-1}|f(x)| < \infty$ and $\int_0^\infty x^{-\sigma}\,\mathrm{d}\mu < \infty$, i.e., the Mellin transforms $\widetilde{f}(s)$ and $\widetilde{\mu}(1-s)$ are defined when $\Re s = \sigma$. Suppose also that the integral $\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \widetilde{f}(s)\widetilde{\mu}(1-s)\,\mathrm{d}s$ converges at least conditionally. Suppose further that $x^\sigma f(x)$ is bounded and that f is μ-a.e. continuous. Then*

$$\int_0^\infty f(x)\,\mathrm{d}\mu(x) = \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \widetilde{f}(s)\widetilde{\mu}(1-s)\,\mathrm{d}s. \tag{10.1}$$

We may apply this lemma to (4.4) and similar formulas, and then obtain an asymptotic expansion from the complex line integral in (10.1) by shifting the line of integration using residue calculus. This gives an alternative method to obtain the asymptotic expansions found above. The new method is perhaps less intuitive that the method used above, and although it does not require Lemma 6.1, it requires some technical arguments and estimates in the complex plane, somewhat similar to the proof of Lemma 6.1. On the other hand, granted these technical details, the method is straightforward and presents the resulting asymptotics in a more structured form as a sum of residues (see for example (10.15)). This is perhaps the main advantage of the method; it is not important for the simple cases studied so far, but it will be essential in Section 11 when we look at the more complicated case of higher moments.

Parseval's formula for Mellin transforms has long been used to derive asymptotic expansions for various integrals and integral transforms in a way similar to our use here, see for example [14] and [21], but this is to our knowledge the first application to combinatorial probability. (Mellin transforms are used in other ways in many combinatorial problems, see for example [9].)

10.1. **Asymptotics of $\mathbb{E}[D_n]$: method 2.** As a warmup, we return to $\mathbb{E}[D_n]$ and give a second proof of Theorem 7.3, using Lemma 10.1 instead of Lemma 6.1. We begin with some simple estimates that will be used repeatedly.

Lemma 10.2. *For $n \ge 2$ and any complex $s = \sigma + \mathrm{i}\tau$, we have*

$$\begin{aligned}\left|\frac{\Gamma(s)\Gamma(n)}{\Gamma(s+n)}\right| &= \frac{\Gamma(n)}{|s(s+1)|\prod_{j=2}^{n-1}|s+j|} \le \frac{\Gamma(n)}{|s(s+1)|\prod_{j=2}^{n-1}|\sigma+j|} \\ &= \frac{\Gamma(\sigma+2)\Gamma(n)}{|s(s+1)|\,\Gamma(\sigma+n)}.\end{aligned} \tag{10.2}$$

Hence:

(i) *For a fixed $n \ge 2$, uniformly for σ in any compact subset of $(-1,\infty)$,*

$$\left|\frac{\Gamma(s)\Gamma(n)}{\Gamma(s+n)}\right| = O\big(|s|^{-2}\big). \tag{10.3}$$

(ii) *For a fixed $\sigma > -1$, uniformly for all n,*

$$\left|\frac{\Gamma(s)\Gamma(n)}{\Gamma(s+n)}\right| = O\big(|s|^{-2}n^{-\sigma}\big). \tag{10.4}$$

Proof. First, (10.2) is elementary, using $\Gamma(z+1) = z\Gamma(z)$. If $\sigma > -1+\delta$ for some $\delta > 0$, then $\sigma \le C(\sigma+1) \le C|s+1|$ for some $C = C(\delta)$, and it follows that $|s| \le C|s+1|$ and thus $|s+1|^{-1} \le C|s|^{-1}$. Hence (10.3) follows from (10.2). Furthermore, for a fixed σ, $\Gamma(n)/\Gamma(n+\sigma) \sim n^{-\sigma}$ as $n \to \infty$, see [19, 5.11.12] (or (7.18)), and thus $\Gamma(n)/\Gamma(n+\sigma) = O\bigl(n^{-\sigma}\bigr)$. Hence, (10.4) too follows from (10.2). ☐

Lemma 10.3. (i) *For $s = \sigma + \mathrm{i}\tau$ with any fixed real $\sigma \notin \{1 - s_i : i \ge 0\}$, we have*

$$|\psi(1-s) - \psi(1)| \ge c \tag{10.5}$$

for some $c = c(\sigma) > 0$.

(ii) *The lower bound* (10.5) *holds uniformly for σ in any compact subset of $\mathbb{R}$ and $|\tau| \ge 1$.*

Proof. The left-hand side of (10.5) is in both cases a continuous function of s that is non-zero on the given sets of s by Lemma 3.1; furthermore it tends to ∞ by (3.6) as $|\tau| \to \infty$ (with σ fixed or in a compact set). Hence (10.5) holds on the given sets of s by a compactness argument. ☐

Second proof of Theorem 7.3. We apply Lemma 10.1 to the measure Υ and the function f_n in (7.1). Note that f_n is continuous everywhere except at 1, and that $\Upsilon\{1\} = 0$ since Υ is concentrated on $(0,1)$ (by (4.1), because each $P_{t,1} < 1$ a.s.); hence f_n is continuous Υ-a.e. as required. Note also that $f_n(x) = O(x)$ on $(0,1)$, so $x^\sigma f_n(x)$ is bounded for $\sigma \ge -1$. The Mellin transform $\widetilde{f}_n(s)$ is given by Lemma 7.1; it is defined for $\Re s > -1$.

Fix $n \ge 2$ and let $s = \sigma + \mathrm{i}\tau$ for some fixed $\sigma \in (-1,0)$. Then (7.4) implies that

$$|\widetilde{f}_n(s)| = \Bigl|\frac{1}{s} + O\Bigl(\frac{1}{|s|^2}\Bigr)\Bigr| \sim \frac{1}{|s|} \sim \frac{1}{|\tau|} \tag{10.6}$$

as $\tau \to \pm\infty$. Furthermore, (4.8) and (3.6) show that, for $|\tau| \ge 2$, say,

$$|\widetilde{\Upsilon}(1-s)| = \bigl|\log(1-s) + O(1)\bigr|^{-1} = \bigl(\log|\tau| + O(1)\bigr)^{-1}. \tag{10.7}$$

Consequently, $|\widetilde{f}_n(s)\widetilde{\Upsilon}(1-s)| \sim 1/(|\tau|\log|\tau|)$ as $\tau \to \pm\infty$, and thus $\widetilde{f}_n(s)\widetilde{\Upsilon}(1-s)$ is *not* absolutely integrable on any line $\Re s = \sigma$. Nevertheless, $\widetilde{f}_n(s)\widetilde{\Upsilon}(1-s)$ is conditionally integrable on the line $\Re s = \sigma$. This can easily be shown, but we postpone this since it will follow from the calculations below. (Conditional integrability is all that we need, but the absence of absolute integrability is a nuisance that complicates the arguments.)

Assuming this conditional integrability, we have verified the conditions of Lemma 10.1, and consequently (10.1) holds, which by (4.4), (4.8), and (7.4) yields, for $-1 < \sigma < 0$,

$$\mathbb{E}[D_n] = \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \widetilde{f}_n(s)\widetilde{\Upsilon}(1-s)\,\mathrm{d}s = \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\widetilde{f}_n(s)}{\psi(1-s)-\psi(1)}\,\mathrm{d}s \tag{10.8}$$

$$= \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{1}{s(\psi(1-s)-\psi(1))}\,\mathrm{d}s - \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{\psi(1-s)-\psi(1)}\,\mathrm{d}s, \tag{10.9}$$

where the second integral in (10.9) (but not the first) is absolutely convergent, by (10.3) and (10.5). To treat the the first integral in (10.9) we split it into three as

$$\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{1}{s}\Bigl(\frac{1}{\psi(1-s)-\psi(1)}-\frac{1}{\log(1-s)}\Bigr)\,\mathrm{d}s \\ +\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\Bigl(\frac{1}{s}+\frac{1}{1-s}\Bigr)\frac{1}{\log(1-s)}\,\mathrm{d}s-\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{1}{1-s}\frac{1}{\log(1-s)}\,\mathrm{d}s. \tag{10.10}$$

The third integral in (10.10) has the primitive function $-\log\log(1-s)$. For $s=\sigma+\mathrm{i}\tau$ with $\sigma<0$, we have $|1-s|\geq 1+\sigma>1$ and thus $\log(1-s)$ lies in the right half-plane. Furthermore, as $\tau\to\pm\infty$, $\log(1-s)=\log|\tau|+O(1)$ and thus $\log\log(1-s)=\log\log|\tau|+o(1)$. It follows that the third integral in (10.10) is conditionally integrable, and that its value is 0. (Note that it is important here that the definition uses the limit of symmetric integrals $\int_{-A}^{A}$.)

We will show that the first two integrals in (10.10) converge absolutely. This means that all terms in (10.10) are well-defined, and that we can sum them to obtain the first integral in (10.9), which in turn means that this integral and the integrals in (10.8) are conditionally convergent, which justifies the use of Lemma 10.1 above.

For the first integral in (10.10), we note first that the integrand is analytic in the domain $\Re s<0$ since $\psi(1-s)-\psi(1)\neq 0$ there by Lemma 3.1, and that in this domain we have by (3.6) $\psi(1-s)=\log(1-s)+O(1)$ and thus, uniformly in any half space $\Re s\leq\delta<0$,

$$\frac{1}{\psi(1-s)-\psi(1)}-\frac{1}{\log(1-s)}=\frac{O(1)}{(\psi(1-s)-\psi(1))\log(1-s)} \\ =O\Bigl(\frac{1}{|\log(1-s)|^2}\Bigr)=O\Bigl(\frac{1}{\log^2|1-s|}\Bigr)=O\Bigl(\frac{1}{\log^2(2+\tau)}\Bigr). \tag{10.11}$$

It follows that the integrand in the first integral can be bounded for $\Re s\leq-\delta$ by

$$\Bigl|\frac{1}{s}\Bigl(\frac{1}{\psi(1-s)-\psi(1)}-\frac{1}{\log(1-s)}\Bigr)\Bigr|\leq\frac{C(\delta)}{(|\sigma|+|\tau|)\log^2(2+\tau)}. \tag{10.12}$$

Hence the integral converges absolutely for every $\sigma<0$. Furthermore, (10.12) implies also that we may shift the line of integration, and thus the value of the integral is the same for all $\sigma<0$; moreover, (10.12) and dominated convergence shows that as $\sigma\to-\infty$, the value converges to 0. Consequently, the first integral in (10.10) vanishes for every $\sigma<0$.

For the second integral in (10.10) we argue similarly, but simpler. In any half space $\Re s\leq\delta<0$, we have $|\log(1-s)|\geq\log|1-s|\geq\log(1+\delta)$ and thus

$$\Bigl(\frac{1}{s}+\frac{1}{1-s}\Bigr)\frac{1}{\log(1-s)}=\frac{1}{s(1-s)}\frac{1}{\log(1-s)}=O\Bigl(\frac{1}{|s(1-s)|}\Bigr)=O\Bigl(\frac{1}{|s|^2}\Bigr). \tag{10.13}$$

It follows again that the integral is absolutely convergent for every $\sigma<0$ and that we may let $\sigma\to-\infty$ and conclude that the integral vanishes.

This completes the proof that the first integral in (10.9) is conditionally convergent, and shows also that it is 0. Hence, (10.8)–(10.9) finally simplifies to

$$\mathbb{E}\,[D_n] = -\frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{\psi(1-s)-\psi(1)}\,\mathrm{d}s, \tag{10.14}$$

valid for any $n \geq 2$ and $\sigma \in (-1, 0)$.

Denote the integrand in (10.14) by $g(s)$; then $g(s)$ is analytic in $\{\Re s > -1\}$ except for poles $\{1 - s_i : i \geq 0\}$ (including $1 - s_0 = 0$). For any $n \geq 2$ and $\sigma \notin \{1 - s_i : i \geq 0\}$ with $\sigma > -1$, (10.3) and (10.5) show that the integral in (10.14) converges absolutely. Moreover, if we assume $|\tau| \geq 1$, then (10.3) and Lemma 10.3(ii) show that the integrand in (10.14) is $O\bigl(|s|^{-2}\bigr)$ uniformly for σ in any compact subset of $(-1, \infty)$, and it follows that we may move the line of integration from a $\sigma \in (-1, 0)$ to a (large) positive σ, picking up $-2\pi\mathrm{i}$ times the residues at the passed poles.

So far we have argued with a fixed n; now we let $n \geq 2$ be arbitrary. Then, for a fixed $\sigma = \Re s > 0$ not in the set $\{1 - s_i : i \geq 0\}$ of poles, (10.4) and (10.5) imply that the integral in (10.14) is $O\bigl(n^{-\sigma}\bigr)$. Consequently, (10.14) implies that $\mathbb{E}\,[D_n]$ has an asymptotic expansion consisting of the residues of $g(s)$ at its poles:

$$\mathbb{E}\,[D_n] = \sum_{i=0}^{N}\operatorname{Res}_{s=1-s_i}\frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{\psi(1-s)-\psi(1)} + O\bigl(n^{-\sigma}\bigr) \tag{10.15}$$

for any fixed $N \geq 0$ and $\sigma < 1 - s_{N+1}$. The first pole $0 = 1 - s_0$ has order 2 (since also $\Gamma(s)$ has a pole there) and a straightforward calculation, using (3.3)–(3.5) and $\psi''(1) = -2\zeta(3)$ (which follows from (3.10) or [19, 5.7.4]) yields

$$\operatorname{Res}_{s=0}\frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{\psi(1-s)-\psi(1)} = \frac{\psi(n)-\psi(1)}{\psi'(1)} - \frac{\psi''(1)}{2\psi'(1)^2} = \frac{h_{n-1}}{\zeta(2)} + \frac{\zeta(3)}{\zeta(2)^2}. \tag{10.16}$$

For any other pole $1 - s_i = 1 + |s_i|$ ($i \geq 1$), $g(s)$ has a pole of order 1, with residue

$$\operatorname{Res}_{s=1-s_i}\frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{\psi(1-s)-\psi(1)} = -\frac{\Gamma(|s_i|+1)\Gamma(n)/\Gamma(n+|s_i|+1)}{\psi'(s_i)}. \tag{10.17}$$

Hence, (10.15) yields (7.13). Finally, (7.14) follows from (7.13) as before, which completes the proof of Theorem 7.3. □

10.2. **The hop-heights L_n: method 2.**

We find it instructive to also use Lemma 10.1 to give another proof of Theorem 8.1.

Second proof of Theorem 8.1. By (8.5), $H_n(x)$ is a polynomial on $[0, 1]$ with $H_n(0) = 0$, and the Mellin transform $\widetilde{H}_n(s)$ exists for $\Re s > -1$; also, $x^\sigma H_n(x)$ is bounded for $\sigma > -1$. Hence, Lemma 10.1 applies to $H_n(x)$ and Υ and any $\sigma \in (-1, 0)$, provided the second integral in (10.1) exists at least conditionally.

We thus obtain, by (8.4), Lemma 10.1, (8.9), and (4.8), for $-1 < \sigma < 0$,

(10.18)
$$\mathbb{E}\,[L_n] = \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \widetilde{H}_n(s)\widetilde{\Upsilon}(1-s)\,\mathrm{d}s = \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{h_{n-1}-\widetilde{f}_n(s)}{s(\psi(1-s)-\psi(1))}\,\mathrm{d}s$$

(10.19)
$$= \frac{h_{n-1}}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{1}{s(\psi(1-s)-\psi(1))}\,\mathrm{d}s - \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\widetilde{f}_n(s)}{s(\psi(1-s)-\psi(1))}\,\mathrm{d}s$$

provided the last two integrals exist at least conditionally. In fact, the first integral in (10.19) is the same as in (10.9), which we have shown converges and equals 0. Furthermore, (10.6) and (10.5) show that for any fixed $\sigma \in (-1, 0)$, the integrand in the second integral in (10.19) is $O\big(|s|^{-2}\big)$ and thus this integral is absolutely convergent. Consequently, (10.18)–(10.19) are justified, and simplify to, using (7.4), and with absolutely convergent integrals,

(10.20)
$$\begin{aligned}\mathbb{E}\,[L_n] &= -\frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\widetilde{f}_n(s)}{s(\psi(1-s)-\psi(1))}\,\mathrm{d}s\\ &= -\frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{1}{s^2(\psi(1-s)-\psi(1))}\,\mathrm{d}s + \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{s(\psi(1-s)-\psi(1))}\,\mathrm{d}s.\end{aligned}$$

In the first integral on the right-hand side, we may again shift $\sigma \to -\infty$; the integrand is analytic for $\Re s < 0$, and $O\big(|s|^{-2}\big)$ is any halfplane $\Re s \le \delta < 0$, and it follows (by dominated convergence) that the limit, and thus the integral, is 0. Consequently, (10.20) simplifies to

$$\mathbb{E}\,[L_n] = \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{s(\psi(1-s)-\psi(1))}\,\mathrm{d}s, \qquad -1 < \sigma < 0. \tag{10.21}$$

In the half-plane $\Re s > -1$, the integrand in (10.21) has poles at 0 and $1-s_i = 1+|s_i|$ for $i \ge 1$; there is a triple pole at 0, and all other poles are simple. As in Section 10.1, we may shift σ to a large positive value; in fact, the integrand in (10.21) is $g(s)/s$, with $g(s)$ as in Section 10.1, so we may just reuse the estimates in Section 10.1. This yields

$$\mathbb{E}\,[L_n] = -\sum_{i=0}^{N} \operatorname{Res}_{s=1-s_i} \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{s(\psi(1-s)-\psi(1))} + O\big(n^{-\sigma}\big) \tag{10.22}$$

for any fixed $N \ge 0$ and $\sigma < 1 - s_{N+1}$. The residue at 0 is by straightforward (but tedious) calculation -1 times

$$\frac{3}{\pi^2}h_{n-1}^2 + \frac{\zeta(3)}{\zeta(2)^2}h_{n-1} + \frac{\zeta(3)^2}{\zeta(2)^3} + \frac{1}{10} - \frac{3}{\pi^2}\psi'(n) \tag{10.23}$$

and the residues at the simple poles are immediate. Hence, (10.22) yields (8.1). □

Note that in the proofs of Theorems 7.3 and 8.1 just given, the significant differences between $\mathbb{E}\,[D_n]$ and $\mathbb{E}\,[L_n]$ is that in the proof of Theorem 7.3 we have a pole at 0 of order 2, while the corresponding function in the proof of Theorem 8.1 has a pole of order 3. This explains the different powers of $\log n$ in the leading term; it also explains why the calculations for $\mathbb{E}\,[L_n]$ in the first proof in Section 8 are somewhat more complicated that for $\mathbb{E}\,[D_n]$ in Section 7.

10.3. **The length Λ_n: method 2.** Lemma 10.1 yields also an alternative proof of Theorem 9.1.

Second proof of Theorem 9.1. This is similar to the proofs above, so we omit some details. Recalling (9.1)–(9.3), we may apply Lemma 10.1 to λ_n and Υ, again for any $n \geq 2$ and $-1 < \sigma < 0$. As in Section 10.1, this leads to (after discarding one conditionally convergent integral that is 0), for $-1 < \sigma < 0$,

$$\mathbb{E}\,\Lambda_n = -\frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\Gamma(s+1)\Gamma(n+1)/\Gamma(n+s)}{(s-1)(\psi(1-s)-\psi(1))}\,\mathrm{d}s. \tag{10.24}$$

We again shift the line of integration to a large σ, noting that in the half-plane $\Re s > -1$, the integrand has poles at $\{1-s_i : i \geq 0\}$, while $s = 1$ is a removable singularity and not a pole since $\psi(1-s)$ has a pole there. Hence we obtain, similarly to (10.15),

$$\mathbb{E}\,[\Lambda_n] = \sum_{i=0}^{N} \operatorname{Res}_{s=1-s_i} \frac{\Gamma(s+1)\Gamma(n+1)/\Gamma(n+s)}{(s-1)(\psi(1-s)-\psi(1))} + O\big(n^{1-\sigma}\big) \tag{10.25}$$

for any fixed $N \geq 0$ and $\sigma < 1 - s_{N+1}$. All poles are simple, so it is straightforward to compute the residues. In particular, the residue at 0 is $n/\psi'(1) = n/\zeta(2)$, and we obtain (9.10). ◻

11. Higher moments of D_n

The results on $\mathbb{E}\,[D_n]$ above may be extended to higher moments. For any integer $k \geq 1$, we have by (4.11)

$$\mathbb{E}\,[D_n^k] = k\int_0^\infty t^{k-1}\mathbb{P}(D_n > t) = k\int_0^\infty t^{k-1}\,\mathbb{E}\big[1-(1-P_{t,1})^{n-1}\big]\,\mathrm{d}t. \tag{11.1}$$

Thus, if we define an infinite measure on $(0,1)$ by

$$\Upsilon_k := k\int_0^\infty t^{k-1}\mathcal{L}(P_{t,1})\,\mathrm{d}t \tag{11.2}$$

(so $\Upsilon_1 = \Upsilon$ defined in (4.1)), then

$$\mathbb{E}\,[D_n^k] = \int_0^1 \big[1-(1-x)^{n-1}\big]\,\mathrm{d}\Upsilon_k(x). \tag{11.3}$$

As in the special case $k = 1$ in (4.8), we obtain the Mellin transform (i.e., moments) of Υ_k from (2.4):

$$\begin{aligned}(11.4)\quad \int_0^1 x^{s-1}\,\mathrm{d}\Upsilon_k(x) &= k\int_0^1 t^{k-1}\int_0^\infty x^{s-1}\,\mathrm{d}\mathcal{L}(P_{t,1})(s)\,\mathrm{d}t = k\int_0^1 t^{k-1}\,\mathbb{E}\,[P_{t,1}^{s-1}]\,\mathrm{d}t \\ &= k\int_0^1 t^{k-1}e^{-t(\psi(s)-\psi(1))}\,\mathrm{d}t \\ &= k!\,\bigl(\psi(s)-\psi(1)\bigr)^{-k}, \qquad \Re s > 1.\end{aligned}$$

Again, the Mellin transform extends to a meromorphic function in the complex plane, with poles at $1 > s_1 > s_2 > \dots$ given by Lemma 3.1.

11.1. **Asymptotics of $\mathbb{E}\,[D_n^k]$.** Fix $k \geq 2$. We apply Lemma 10.1 to Υ_k and the function $f_n(x)$ in (7.1) as for the case $k = 1$ in Section 10.1. If $s = \sigma + \mathrm{i}\tau$ with a fixed $\sigma \in (-1, 0)$, then (11.4) and (10.7) show that

$$(11.5)\qquad |\widetilde{\Upsilon}_k(1-s)| = k!\,|\widetilde{\Upsilon}(1-s)|^k = \bigl(\log|\tau| + O(1)\bigr)^{-k},.$$

which combined with (10.6) shows that $\widetilde{f}_n(s)\widetilde{\Upsilon}_k(1-s)$ is absolutely integrable on the line $\Re s = \sigma$. (This simplifies the argument needed for $k = 1$ in Section 10.1.) The other conditions of Lemma 10.1 are satisfied as in Section 10.1, in particular, f_n is continuous Υ_k-a.e. since $\Upsilon_k\{1\} = 0$. Consequently (10.1) holds, which by (11.3), (7.1), and (11.4) yields

$$\begin{aligned}(11.6)\qquad \mathbb{E}\,[D_n^k] &= \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\widetilde{f}_n(s)\widetilde{\Upsilon}_k(1-s)\,\mathrm{d}s \\ &= \frac{k!}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{\widetilde{f}_n(s)}{(\psi(1-s)-\psi(1))^k}\,\mathrm{d}s, \qquad -1<\sigma<0.\end{aligned}$$

We split $\widetilde{f}_n(s)$ into two parts according to (7.4); thus (11.6) yields

$$\begin{aligned}(11.7)\qquad \mathbb{E}\,[D_n^k] =& \frac{k!}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{1}{s(\psi(1-s)-\psi(1))^k}\,\mathrm{d}s \\ &- \frac{k!}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{(\psi(1-s)-\psi(1))^k}\,\mathrm{d}s, \qquad -1<\sigma<0.\end{aligned}$$

In the first integral in (11.7), the integrand is analytic in the half-plane $\{\Re s < 0\}$. Furthermore, in this half-plane $\psi(1-s) = \log(1-s) + O(1) = \log(|s|+1) + O(1)$, see again (3.6). It follows first that we may move the line of integration to any $\sigma < 0$, and then that as $\sigma \to -\infty$, the integral tends to 0 by dominated convergence. (Again, the case $k \geq 2$ is simpler that $k = 1$ here, although the conclusion is the same.) Consequently, the first integral in (11.7) is 0, and thus we have, as for the case $k = 1$ in (10.14),

$$(11.8)\qquad \mathbb{E}\,[D_n^k] = -\frac{k!}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{(\psi(1-s)-\psi(1))^k}\,\mathrm{d}s,$$

for any $n \geq 2$ and $\sigma \in (-1, 0)$.

We then argue as in Section 10.1, and move the line of integration in (11.8) to a (large) positive σ; this is again justified by the estimates in Lemmas 10.2 and 10.3. The integrand, $g_k(s)$ say, is again analytic in $\{\Re s > -1\}$ except for poles at 0 and $\{1 - s_i = |s_i| + 1 : i \geq 1\}$, and we obtain again an asymptotic expansion consisting of the residues of $g_k(s)$ at its poles:

$$\mathbb{E}\,[D_n^k] = k! \sum_{i=0}^{N} \operatorname{Res}_{s=1-s_i} \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{(\psi(1-s)-\psi(1))^k} + O\bigl(n^{-\sigma}\bigr) \tag{11.9}$$

for any fixed $N \geq 0$ and $\sigma < 1 - s_{N+1}$. The first pole $0 = 1 - s_0$ has order $k+1$ so the residue there can be obtained by computing the coefficients of the singular part of the Laurent expansion of $\Gamma(s)/\bigl((\psi(1-s)-\psi(1))^k\bigr)$ and the $k+1$ first Taylor coefficients of $\Gamma(n)/\Gamma(n+s)$ at $s = 0$; it follows from (3.3) that the latter coefficients are by polynomials in $\psi(n)$ and the derivatives $\psi'(n), \dots, \psi^{(k-1)}(n)$; these may by (3.6) all be expanded further into asymptotic expansions involving $\log n$ (for $\psi(n)$ only) and powers of n^{-1}. The leading term will be a constant times $\log^k n$, in fact, it is easily seen to be (including the factor $k!$ in (11.9))

$$\psi'(1)^{-k} \log^k n = \bigl(\tfrac{6}{\pi^2} \log n\bigr)^k. \tag{11.10}$$

For any other pole $1 - s_i = 1 + |s_i|$ $(i \geq 1)$, $g_k(s)$ has a pole of order k, and the contribution from this pole to (11.9) is given by some combination (with computable coefficients) of the k first Taylor coefficients of $\Gamma(n)/\Gamma(n+s)$ at $s = 1 + |s_i|$. The dominant term will be a (nonzero) constant times $\psi(n)^{k-1}\Gamma(n)/\Gamma(n+1+|s_i|) \sim (\log n)^{k-1} n^{-|s_i|-1}$, so the entire residue is of this order. It follows that the expansion (11.9) is an asymptotic expansion of the type defined in (3.1)–(3.2) with terms of successively smaller order; hence we may write (11.9) as

$$\mathbb{E}\,[D_n^k] \sim k! \sum_{i=0}^{\infty} \operatorname{Res}_{s=1-s_i} \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{(\psi(1-s)-\psi(1))^k}, \tag{11.11}$$

where the residues may calculated and expanded as above, leading to an asymptotic expansion containing terms that are constants times $(\log n)^k$ (the leading term), $(\log n)^j n^{-\ell}$ $(0 \leq j \leq k-1,\ \ell \geq 0)$, and $(\log n)^j n^{-|s_i|-\ell}$ $(i \geq 1,\ 0 \leq j \leq k-1,\ \ell \geq 1)$; the coefficients may be computed as sketched above. In particular, we obtain by collecting terms:

Theorem 11.1. *For any fixed integer $k \geq 1$, we have as $n \to \infty$ the asymptotic expansion* (11.11)*, and in particular*

$$\mathbb{E}\,[D_n^k] = p_k(\log n) + O\Bigl(\frac{\log^{k-1} n}{n}\Bigr) \tag{11.12}$$

for some (computable) polynomial p_k of degree k, with leading term (11.10)*.*

For example, for $k = 2$, the residue at 0 in (11.11) is, including the factor $2!$ and recalling (3.5)

$$\frac{h_{n-1}^2}{\zeta(2)^2} + \frac{4\zeta(3)}{\zeta(2)^3} h_{n-1} + \frac{6\zeta(3)^2}{\zeta(2)^4} - \frac{18}{5\pi^2} - \frac{1}{\zeta(2)^2}\psi'(n). \tag{11.13}$$

Hence, (11.11) yields, using (3.8),

$$\begin{aligned}\mathbb{E}\,[D_n^2] &= \frac{1}{\zeta(2)^2}h_{n-1}^2 + \frac{4\zeta(3)}{\zeta(2)^3}h_{n-1} + \frac{6\zeta(3)^2}{\zeta(2)^4} - \frac{18}{5\pi^2} + O\Big(\frac{1}{n}\Big) \\ &= \frac{1}{\zeta(2)^2}\log^2 n + \Big(\frac{2\gamma}{\zeta(2)^2} + \frac{4\zeta(3)}{\zeta(2)^3}\Big)\log n \\ &\qquad + \frac{\gamma^2}{\zeta(2)^2} + \frac{4\zeta(3)}{\zeta(2)^3}\gamma + \frac{6\zeta(3)^2}{\zeta(2)^4} - \frac{18}{5\pi^2} + O\Big(\frac{\log n}{n}\Big).\end{aligned} \tag{11.14}$$

Combining (11.14) and (7.13), we obtain the following result; the leading term was found by a different method in [5, Theorem 1.1].

Theorem 11.2. *The variance of D_n is, as $n\to\infty$,*

$$\begin{aligned}\operatorname{var}[D_n] &= \frac{2\zeta(3)}{\zeta(2)^3}h_{n-1} + \frac{5\zeta(3)^2}{\zeta(2)^4} - \frac{18}{5\pi^2} + O\Big(\frac{\log n}{n}\Big) \\ &= \frac{2\zeta(3)}{\zeta(2)^3}\log n + \frac{2\zeta(3)}{\zeta(2)^3}\gamma + \frac{5\zeta(3)^2}{\zeta(2)^4} - \frac{18}{5\pi^2} + O\Big(\frac{\log n}{n}\Big).\end{aligned} \tag{11.15}$$

11.2. **Asymptotics of Υ_k.** The argument above did not use any properties of Υ_k except its Mellin transform and $\Upsilon_k\{1\} = 0$. For completeness, we also consider results analoguous to Lemma 6.1.

We may invert the Mellin transform by the method in Section 6 and show that Υ_k is absolutely continuous and obtain an asymptotic expansion of its density $\nu_k(x)$ as $x \searrow 0$. However, now each pole has order k, which makes the Mellin inversion a bit more complicated.

Consider for simplicity first the case $k = 2$. Then (11.4) shows that Υ_2 has the Mellin transform $2(\psi(s) - \psi(1))^{-2}$, which has double poles at 1 and every s_i. We argue as in Section 6 (to which we refer for omitted details) and consider the finite measure ν on $(0,1)$ defined by $\mathrm{d}\nu(x) = x\,\mathrm{d}\Upsilon_2(x)$; this shifts the Mellin transform to

$$\widetilde{\nu}(s) = \widetilde{\Upsilon_2}(s+1) = \frac{2}{(\psi(s+1) - \psi(1))^2}, \tag{11.16}$$

which has a double pole at 0 with singular part

$$\frac{2}{\zeta(2)^2}s^{-2} + \frac{4\zeta(3)}{\zeta(2)^3}s^{-1}. \tag{11.17}$$

We therefore define ν_0 as the measure on $(0,1)$ with density

$$h_0(x) := \frac{2}{\zeta(2)^2}(-\log x) + \frac{4\zeta(3)}{\zeta(2)^3}. \tag{11.18}$$

Then ν_0 has Mellin transform exactly (11.17) (for $\Re s > 0$), and thus the signed measure $\nu_\Delta := \nu - \nu_0$ has a Mellin transform which extends to a meromorphic function in $\mathbb{C}$ with poles only at $\{s_i - 1 : i \geq 1\}$ (all double). Again, the Mellin transform is not integrable on any vertical lines, so we use again the trick of considering $\widetilde{\nu_\Delta}'(s)$, which is the Mellin transform of $\log(x)\,\mathrm{d}\nu_\Delta(x)$, and which is integrable on any vertical line not containing a pole s_i. Thus the Mellin inversion formula (6.17) applies. By, as in Section 6, shifting the line of integration and undoing the modifications

$\Upsilon_2 \to \nu \to \nu_\Delta \to (\log x)\,\mathrm{d}\nu_\Delta$ used in the argument, it follows that Υ_2 is absolutely continuous with a density

$$\begin{aligned}(11.19)\quad \upsilon_2(x) := \frac{\mathrm{d}\Upsilon_2}{\mathrm{d}x} &= \frac{1}{x}h_0(x) + \frac{1}{x\log x}O\big(x^{-s_1+1-\varepsilon}\big)\\ &= \frac{2}{\zeta(2)^2}\cdot\frac{-\log x}{x} + \frac{4\zeta(3)}{\zeta(2)^3}\cdot\frac{1}{x} + O\big(x^{|s_1|-\varepsilon}|\log x|^{-1}\big), \qquad 0<x<1,\end{aligned}$$

for any $\varepsilon > 0$. We may improve this to a full asymptotic expansion as in (6.23) by caculating contributions from further poles; note that we now get for each pole s_i two terms, with x^{-s_i} and $(\log x)x^{-s_i}$.

We now may use this and (11.3) to obtain the asymptotical expansion (11.14) for $\mathbb{E}\,[D_n^2]$, but the calculations are more complicated than in Section 7 since we now have more terms, and we leave the details to the interested reader.

From the main term in (11.19) we obtain, using both (7.3) and the result of differentiating it,

$$\begin{aligned}(11.20)\qquad \int_0^1 \big[1-(1-x)^{n-1}\big]\Big(\frac{2}{\zeta(2)^2}\cdot\frac{-\log x}{x} + \frac{4\zeta(3)}{\zeta(2)^3}\cdot\frac{1}{x}\Big)\,\mathrm{d}x\\ = -\frac{2}{\zeta(2)^2}\widetilde{f}_n'(0) + \frac{4\zeta(3)}{\zeta(2)^3}\widetilde{f}_n(0).\end{aligned}$$

For the remainder term in (11.19) we have essentially the same estimate as in (7.8) (with a different constant).

We have computed $\widetilde{f}_n(0) = h_{n-1}$ in (7.6), and $-\widetilde{f}_n'(0)$ in (8.12). Since $\psi'(n) = O(1/n)$, as a consequence of (3.6), we find from (11.3) and (11.20)

$$(11.21)\qquad \mathbb{E}\,[D_n^2] = \frac{1}{\zeta(2)^2}h_{n-1}^2 + \frac{4\zeta(3)}{\zeta(2)^3}h_{n-1} + O(1),$$

as already shown (with more precision) in (11.14). We may also obtain further terms, and a full asymptotic expansion, but the method of Section 11.1 seems preferable.

We may treat moments of any order $k \geq 2$ in the same way, in principle with full asymptotic expansions. In general, the Mellin transform (11.4) has poles of order k, and the main term (for our purposes at least, i.e., for small x) of the density $\upsilon_k(x)$ of Υ_k will be of the form $q_k(\log x)/x$, where q_k is a polynomial of degree $k-1$; the leading term of $\upsilon_k(x)$ is, more precisely, $k\zeta(2)^{-k}(-\log x)^{k-1}/x$. Substituting this in (11.3) and using (7.3)–(7.4) as above yields, unsurprisingly, the main term

$$(11.22)\qquad \mathbb{E}\,[D_n^k] \sim k\zeta(2)^{-k}\Big(-\frac{\mathrm{d}}{\mathrm{d}s}\Big)^{k-1}\widetilde{f}_n(s)\Big|_{s=0} \sim \Big(\frac{h_{n-1}}{\zeta(2)}\Big)^k \sim \Big(\frac{6}{\pi^2}\log n\Big)^k.$$

Further terms can be obtained in a straightforward way, but again the method in Section 11.1 seems preferable.

12. Moment generating function of D_n

Finally, we consider the moment generating function $\mathbb{E}\,[e^{zD_n}]$ of D_n. Note first that conditioned on DTCS(n), i.e., on the structure of the tree but forgetting the edge lengths, the height D_n is a sum of a finite number of exponential random variables, each with expectation $1/h_{m-1} \le 1$ for some $m \ge 2$; it follows that the moment generating function $\mathbb{E}\,[e^{zD_n}]$ exists for $\Re z < 1$, and thus is an analytic function there.

Remark 12.1. The same argument shows that $\mathbb{E}\,[e^{D_n}] = \infty$ for every $n \ge 2$, since with positive probability leaf 1 belongs to a clade of size $m = 2$; hence the moment generating function exists if and only if $\Re z < 1$.

We will derive our results using the (exact) formula (11.8) for moments. We will need a bound for the denominator there, more precise than Lemma 10.3.

Lemma 12.2. *Let $s = \sigma + \mathrm{i}\tau$ with $\sigma > 1$. Then*

$$|\psi(s) - \psi(1)| \ge \Re\bigl(\psi(s) - \psi(1)\bigr) \ge \psi(\sigma) - \psi(1) > 0. \tag{12.1}$$

Proof. The first inequality is trivial, and the last follows since (3.10) implies that $\psi'(z) > 0$ for $z > 0$. Finally, (3.10) also implies that if $\sigma > 0$ and $\tau > 0$, then $\Im\psi'(\sigma + \mathrm{i}\tau) < 0$, and thus $\Re\psi(\sigma + \mathrm{i}\tau)$ increases as τ grows from 0 to ∞; the case $\tau \le 0$ follows by symmetry. ☐

Lemma 12.3. *Let $n \ge 2$ and $-1 < \sigma < 0$. Then, for every complex z with $\Re z < \psi(1-\sigma) - \psi(1)$,*

$$\mathbb{E}\,[e^{zD_n}] = 1 - \frac{z}{2\pi\mathrm{i}} \int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{\psi(1-s) - \psi(1) - z}\,\mathrm{d}s, \tag{12.2}$$

where the integral is absolutely convergent.

Proof. Note first that $1 - \sigma > 1$, so $b := \psi(1-\sigma) - \psi(1) > 0$ by Lemma 12.2, which furthermore shows that $\Re\bigl(\psi(1-s) - \psi(1)\bigr) \ge b$ for $s = \sigma + \mathrm{i}\tau$. Hence, the denominator in the integral in (12.2) is bounded away from 0 when $\Re z < b$, uniformly for z in any compact subset. It follows, using (10.3), that the integral in (12.2) converges absolutely and defines an analytic function of z in the half-plane $\Re z < b$.

Note also that $1 - \sigma < 2$, and thus $b < \psi(2) - \psi(1) = 1$; hence the half-plane $\Re z < b$ is contained in the half-plane $\Re z < 1$ where we know that $\mathbb{E}\,[e^{zD_n}]$ exists and is analytic. By analytic continuation, it thus suffices to show (12.2) for small $|z|$, and we will in the rest of the proof assume $|z| < b$. In particular, $|z| < 1$, and thus $\mathbb{E}\,[e^{|z|D_n}] < \infty$. Consequently we have by (11.8), which holds also for $k = 1$ by (10.14),

$$\mathbb{E}\,[e^{zD_n}] = 1 + \sum_{k=1}^{\infty} \frac{z^k}{k!}\,\mathbb{E}\,[D_n^k] = 1 - \frac{1}{2\pi\mathrm{i}} \sum_{k=1}^{\infty} \int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} z^k \frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{(\psi(1-s) - \psi(1))^k}. \tag{12.3}$$

We may interchange the order of summation and integration by Fubini's theorem, which is justified by (10.3) together with Lemma 12.2 which gives $|\psi(1-s) - \psi(1)| \ge b > |z|$; then (12.2) follows by summing the geometric series. ☐

The next step is to shift the line of integration to the right. To do this in general would require a study of the roots of $\psi(s) - \psi(1) = z$ in the complex plane. We consider for simplicity only the case of real z; extensions to complex z are left to the reader.

12.1. **Real z.** Consider the equation

$$\psi(1+x)-\psi(1)=z \tag{12.4}$$

for a real z. By Lemma 3.1, and using the notation there, the roots are

$$\rho_i(z):=s_i\bigl(z+\psi(1)\bigr)-1, \qquad i=0,1,2,\dots, \tag{12.5}$$

with $\rho_0(z)\in(-1,\infty)$ and $\rho_i(z)\in(-(i+1),-i)$ for $i\geq 1$. We are mainly interested in $\rho(z):=\rho_0(z)$, the largest root; thus $\rho(z)$ is the unique real number in $(-1,\infty)$ satisfying

$$\psi\bigl(1+\rho(z)\bigr)-\psi(1)=z. \tag{12.6}$$

The function $\rho:(-\infty,\infty)\to(-1,\infty)$ is strictly increasing and continuous, in fact analytic (since ψ is on $(0,\infty)$), and (12.6) shows that $\rho(0)=0$. Furthermore, since (3.5) yields $\psi(2)-\psi(1)=1$, (12.6) also shows $\rho(1)=1$. Hence, ρ is a bijection $(-\infty,1)\to(-1,1)$.

Fix a real $z\in(-\infty,1)$. Then, as just said, $\rho(z)\in(-1,1)$. The zeroes of the denominator in (12.2) are $s=-\rho(z)\in(-1,1)$ and $-\rho_1(z)<-\rho_2(z)<\dots$, with $-\rho_1(z)\in(1,2)$. Write for convenience $\rho:=\rho(z)$, and take $\sigma\in(-1,\min(\rho,0))$. Then

$$\psi(1-\sigma)-\psi(1)>\psi\bigl(1-\rho\bigr)-\psi(1)=z, \tag{12.7}$$

and thus Lemma 12.3 applies.

We now shift the line of integration in (12.2) to some $\sigma\in(1,-\rho_1(z))$; this is easily justified using (10.3) and noting that the proof of Lemma 10.3(ii) also shows, more generally, that $|\psi(1-s)-\psi(1)-z|\geq c>0$ when $|\tau|\geq 1$ for any $z\in\mathbb{R}$. We pass two poles of the integrand, at $s=0$ and $s=-\rho$, and we pick up $-2\pi\mathrm{i}$ times the residues there. We assume that $z\neq 0$, since the case $z=0$ is trivial; then these two poles are distinct and both are simple. The residue at 0 of the integrand in (12.2) is simply $-1/z$, so the contribution there is -1, which cancels the constant 1 in (12.2). The main term comes from the residue at $-\rho$, which is $-\Gamma(-\rho)\Gamma(n)/[\Gamma(n-\rho)\psi'(1+\rho)]$. Hence we obtain, for any $1<\sigma<-\rho_1(z)=1+|s_1(z-\gamma)|$,

(12.8)
$$\mathbb{E}\,[e^{zD_n}]=-z\frac{\Gamma(-\rho(z))\Gamma(n)}{\psi'(1+\rho(z))\Gamma(n-\rho(z))}-\frac{z}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty}\frac{\Gamma(s)\Gamma(n)/\Gamma(n+s)}{\psi(1-s)-\psi(1)-z}\,\mathrm{d}s.$$

Since $z<1$, we have $\rho_1(z)<\rho_1(1)$, and thus we may here always choose σ as

$$\sigma_*:=-\rho_1(1)=1-s_1(1+\psi(1))=1+|s_1(\psi(2))|\doteq 1.457. \tag{12.9}$$

We note also that as $z\to 0$, we have $\rho(z)\to 0$ and

$$-z\Gamma(-\rho(z))=\frac{z}{\rho(z)}\Gamma(1-\rho(z))\to\frac{1}{\rho'(0)}=\psi'(1), \tag{12.10}$$

since $\psi'(1)\rho'(0)=1$ follows by differentiation of (12.6). We thus interpret $-z\Gamma(-\rho(-z))=\psi'(1)$ for $z=0$ and note that then (12.8) holds trivially for $z=0$ too.

This leads to the following result.

Theorem 12.4. *For any real $z < 1$,* (12.8) *holds for all $n \geq 2$ with $\sigma = \sigma_*$ given by* (12.9)*. Hence,*

$$\mathbb{E}\,[e^{zD_n}] = \frac{-z\Gamma(-\rho(z))}{\psi'(1+\rho(z))}\frac{\Gamma(n)}{\Gamma(n-\rho(z))} + O\bigl(n^{-\sigma_*}\bigr) \tag{12.11}$$

and

$$\mathbb{E}\,[e^{zD_n}] = \frac{-z\Gamma(-\rho(z))}{\psi'(1+\rho(z))}n^{\rho(z)} \cdot \bigl(1 + O\bigl(n^{-\min(1,\sigma_*+\rho(z))}\bigr)\bigr). \tag{12.12}$$

Furthermore, (12.11) *holds uniformly for $z < 1-\delta$ for any $\delta > 0$, and* (12.12) *holds uniformly for z in a compact subset of* $(-\infty, 1)$*.*

Proof. We have already shown that (12.8) holds with $\sigma = \sigma_*$. Furthermore, by (12.9) and (12.4)–(12.5),

$$\psi(1-\sigma_*) - \psi(1) = \psi(1+\rho_1(1)) - \psi(1) = 1. \tag{12.13}$$

Hence, if $s = \sigma_* + \mathrm{i}\tau$, then the denominator $\psi(1-s)-\psi(1)-z$ in the integral in (12.8) is $1-z > 0$ when $\tau = 0$, and as seen in the proof of Lemma 3.1, it is non-real for all $\tau \neq 0$; furthermore, $\Re\bigl(\psi(1-s)-\psi(1)-z\bigr) \to +\infty$ as $\tau \to \pm\infty$ by (3.6). It follows, by continuity and compactness, that $\bigl|\psi(1-s)-\psi(1)-z\bigr|$ is bounded below by some $c(\delta) > 0$ for all such s and $z \leq 1-\delta$; moreover, $\Re\bigl(\psi(1-s)-\psi(1)-z\bigr) \geq -C-z \geq |z|/2$ if $z \leq -2C$. It follows, using (10.3) and [19, 5.11.12], that the last term in (12.8) is $O\bigl(\Gamma(n)/\Gamma(n+\sigma_*)\bigr) = O\bigl(n^{-\sigma_*}\bigr)$ uniformly for $z \leq 1-\delta$, which proves (12.11).

The mean value theorem yields, using (3.3) and (3.6), uniformly for $|\rho| \leq 1$,

$$\log\Gamma(n) - \log\Gamma(n-\rho) = \psi\bigl(n+O(1)\bigr)\rho = \rho\log n + O\bigl(n^{-1}\bigr). \tag{12.14}$$

Thus (12.12) follows from (12.11), recalling that $|\rho(z)| < 1$ for all $z < 1$. □

Note that $\sigma_* > 1 > -\rho(z)$, so the exponent in the error term in (12.12) is always negative, and in fact less that $1-\sigma_* = s_2(\psi(2)) \doteq -0.457$. Note also that the proof shows that for any fixed z, the exponent may be improved.

In the following two subsections we give some consequences of Theorem 12.4.

12.2. **A central limit theorem.** As a corollary of Theorem 12.4, we obtain a new proof of the following CLT. As mentioned in Section 1.3, this has been proved in [5, Theorem 1.7] by analysing a recursion for the moment generating function; other proofs by different methods are given in [2], [16], and [18].

Theorem 12.5.

$$\frac{D_n - \mu\log n}{\sqrt{\log n}} \xrightarrow{\mathrm{d}} \mathrm{Normal}(0,\sigma^2) \quad \textit{as } n \to \infty \tag{12.15}$$

where

$$\mu := 1/\zeta(2) = 6/\pi^2 \doteq 0.6079; \quad \sigma^2 := 2\zeta(3)/\zeta(2)^3 \doteq 0.5401. \tag{12.16}$$

Proof. Let $z = z_n$ be any sequence of real numbers with $z_n \to 0$. Then (12.12) and (12.10) yield, as $n \to \infty$,

$$\mathbb{E}\,[e^{z_nD_n}] = n^{\rho(z_n)} \cdot \bigl(1+o(1)\bigr) = e^{\rho(z_n)\log n + o(1)}. \tag{12.17}$$

We have $\rho(0) = 0$, and the implicit function theorem shows that ρ is analytic at 0 (and everywhere). Hence,

$$\rho(z_n) = z_n \rho'(0) + \tfrac{1}{2} z_n^2 \bigl(\rho''(0) + o(1)\bigr). \tag{12.18}$$

Fix $t \in \mathbb{R}$ and let $z_n := t/\sqrt{\log n}$ (for $n \geq 2$). Then (12.17)–(12.18) yield

$$\begin{aligned} \mathbb{E}\, e^{t(D_n - \rho'(0)\log n)/\sqrt{\log n}} &= \mathbb{E}\, e^{z_n(D_n - \rho'(0)\log n)} \\ &= \exp\bigl(\tfrac{1}{2} z_n^2(\rho''(0) + o(1))\log n + o(1)\bigr) \\ &\to e^{\frac{1}{2}t^2\rho''(0)} \end{aligned} \tag{12.19}$$

as $n \to \infty$. Since this holds for every real t, we thus have the CLT (12.15) with $\mu = \rho'(0)$ and $\sigma^2 = \rho''(0)$.

Finally, differentiating (12.6) yields $\rho'(z) = 1/\psi'(1+\rho(z))$ and in particular $\rho'(0) = 1/\psi'(1) = 1/\zeta(2)$ as already noted in (12.10), and differentiating again yields

$$\rho''(0) = -\frac{\psi''(1)}{\psi'(1)^3} = \frac{2\zeta(3)}{\zeta(2)^3}, \tag{12.20}$$

which completes the proof. □

12.3. **Large deviations.** As another corollary of Theorem 12.4, we obtain large deviation results by the Gärtner–Ellis theorem. We follow [7, Section 2.3], but note that n there is replaced by the "speed" $\log n$ (see [7, Remark (a) p. 44]). Thus we define (for $n \geq 2$)

$$Z_n := \frac{1}{\log n} D_n \tag{12.21}$$

and note that (12.12) and Remark 12.1 yield as $n \to \infty$, for any fixed $\lambda \in \mathbb{R}$,

$$\frac{1}{\log n} \log \mathbb{E}\left[e^{(\log n)\lambda Z_n}\right] = \frac{1}{\log n} \log \mathbb{E}\, e^{\lambda D_n} \to \Lambda(\lambda) := \begin{cases} \rho(\lambda), & \lambda < 1, \\ +\infty, & \lambda \geq 1. \end{cases} \tag{12.22}$$

The Fenchel–Legendre transform Λ^* of Λ [7, Definition 2.2.2] is defined by

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}} \bigl\{x\lambda - \Lambda(\lambda)\bigr\} = \sup_{-\infty < \lambda < 1} \bigl\{x\lambda - \rho(\lambda)\bigr\}. \tag{12.23}$$

Recall that $z \mapsto \rho(z)$ is a bijection of $(-\infty, 1)$ onto $(-1, 1)$, which by (12.6) is the inverse function of

$$\rho \mapsto g(\rho) := \psi(1+\rho) - \psi(1). \tag{12.24}$$

Hence, if we define

$$h(x, \rho) := x g(\rho) - \rho, \tag{12.25}$$

then (12.23) yields

$$\Lambda^*(x) = \sup_{-1<\rho<1} \{x g(\rho) - \rho\} = \sup_{-1<\rho<1} h(x,\rho) = \sup_{-1<\rho\leq 1} h(x,\rho) \tag{12.26}$$

where the last equality holds by continuity. We have

$$\frac{\partial}{\partial \rho} h(x,\rho) = x\psi'(1+\rho) - 1 \tag{12.27}$$

and for any fixed $x > 0$, this is by (3.10) strictly decreasing from $+\infty$ to $x\psi'(2)-1 = x(\zeta(2)-1)-1$ as ρ grows from -1 to 1. Hence, for $x > 0$, $h(x,\rho)$ is a concave function of ρ, and the supremum in (12.26) is attained at a unique $\rho_*(x) \in (-1,1]$ given by

$$\begin{cases} \psi'(1+\rho_*(x)) = 1/x, & 0 < x \le x_1 := (\zeta(2)-1)^{-1}, \\ \rho_*(x) = 1, & x \ge x_1. \end{cases} \tag{12.28}$$

For $x \le 0$, $h(x,\rho)$ is a decreasing function of ρ, and thus the supremum in (12.26) is obtained by letting $\rho \to -1$, and thus $g(\rho) \to -\infty$. Thus we obtain, combining the cases,

$$\Lambda^*(x) = \begin{cases} +\infty, & x < 0, \\ 1, & x = 0, \\ xg(\rho_*(x)) - \rho_*(x), & x > 0. \end{cases} \tag{12.29}$$

In particular, for $x \ge x_1$, we have by (12.28) and (12.29)

$$\Lambda^*(x) = xg(1) - 1 = x - 1. \tag{12.30}$$

It follows from (12.23) or (12.26) that Λ^* is convex and that it is continuous on $[0,\infty)$. Taking the derivative in (12.29) yields, for $x > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\Lambda^*(x) = g(\rho_*(x)) + \bigl(xg'(\rho_*(x)) - 1\bigr)\rho_*'(x) = g(\rho_*(x)), \tag{12.31}$$

since (12.28) implies both $xg'(\rho_*(x))-1 = 0$ for $0 < x \le x_1$ and $\rho_*'(x) = 0$ for $x \ge x_1$. (In particular, we see that $\Lambda^*(x)$ is continuously differentiable also at $x = x_1$.) The derivative is thus strictly increasing for $0 < x \le x_1$, so Λ^* is strictly convex on $[0,x_1]$, while $\Lambda^*(x)$ is linear for $x \ge x_1$ as shown already in (12.30). Hence, in the terminology of [7, Definition 2.3.3], $y \in \mathbb{R}$ is an *exposed point* of Λ^* if $0 < y < x_1$.

We note that (12.31) yields $\frac{\mathrm{d}}{\mathrm{d}x}\Lambda^*(x) = 0$ when $\rho_*(x) = 0$ (since $g(0) = 0$), which by (12.28) holds if and only if $x = x_0 := \psi'(1)^{-1} = 6/\pi^2$; furthermore, (12.29) shows that $\Lambda^*(x_0) = 0$. Since Λ^* is convex, it thus attains its minimum at x_0, and the minimum is 0 (as it is has to be).

The Gärtner–Ellis theorem [7, Theorem 2.3.6] now implies Theorem 1.6, restated here.

Theorem 12.6. *As $n \to \infty$, we have:*

$$\mathbb{P}(D_n < x\log n) = n^{-\Lambda^*(x)+o(1)}, \qquad \text{if} \quad 0 < x \le x_0, \tag{12.32}$$

$$\mathbb{P}(D_n > x\log n) = n^{-\Lambda^*(x)+o(1)}, \qquad \text{if} \quad x_0 \le x < x_1, \tag{12.33}$$

$$\mathbb{P}(D_n > x\log n) \le n^{-\Lambda^*(x)+o(1)}, \qquad \text{if} \quad x \ge x_1. \tag{12.34}$$

Theorem 12.6 improves estimates for the upper tail in [5, Theorem 1.4]. As a sanity check we note that (12.32) and (12.33) imply that D_n is concentrated at

$x_0 \log n = \frac{6}{\pi^2} \log n$, and in particular $D_n / \log n \xrightarrow{\mathrm{p}} \frac{6}{\pi^2}$, which is proved in [5], and also follows from Theorem 7.3 and (11.15), or from Theorem 12.5.

Remark 12.7. The results in this section are based on (12.2) which is obtained by summing the corresponding results for the moments $\mathbb{E}[D_n^k]$. For real z, an alternative is to define the (signed) measure

$$\Xi_z := z \int_0^\infty e^{zt} \mathcal{L}(P_{t,1})\,\mathrm{d}t. \tag{12.35}$$

(This is an infinite positive measure for $z > 0$ and a finite negative measure for $z < 0$.) Then

$$\mathbb{E}\, e^{zD_n} = 1 + \int_0^\infty z e^{zt} \mathbb{P}(D_n > t)\,\mathrm{d}t \tag{12.36}$$

and thus (4.11) yields

$$\mathbb{E}\, e^{zD_n} = 1 + \int_0^1 \bigl[1 - (1-x)^{n-1}\bigr]\,\mathrm{d}\Xi_z(x). \tag{12.37}$$

Furthermore, we obtain from (12.35) and (2.4) the Mellin transform, for $\Re s = \sigma > 1$ and $\psi(\sigma) > \psi(1) + z$,

$$\int_0^1 x^{s-1}\,\mathrm{d}\Xi_z(x) = z \int_0^\infty e^{t(z-\psi(s)+\psi(1))}\,\mathrm{d}t = \frac{z}{\psi(s) - \psi(1) - z}. \tag{12.38}$$

For $z < 1$ we may then by the methods above obtain (12.2) and (12.8).

For our purposes, we prefer the method above, but we mention the alternative since it might have other uses.

13. Final remarks

As mentioned in the introduction, this article is part of a broad project investigating different aspects of the random tree model. The document [2] is intended to maintain a current overview of the project, summarizing known results, open problems and heuristics, and indicating the range of proof techniques. Let us mention just two aspects related to this article.

1. The definition of D_n involves two levels of randomness: the realization of the tree and the distribution of leaf heights within that realization. A start at quantifying this feature is a "lack of correlation" result in [5, Theorem 1.6].

2. Heuristics for the height of CTCS(n), that is the maximum leaf height in a realization, are discussed in [2]. This problem seems surprisingly subtle: the naive guess based on the Normal approximation (Theorem 1.5) is definitely incorrect.

Acknowledgments

Thanks to Boris Pittel for extensive interactions regarding this project. Thanks to Bénédicte Haas for his careful explanation of how our setting fits into the general theory of exchangeable random partitions.

Appendix A. Proof of Lemma 10.1

We repeat the statement of Lemma 10.1.

Lemma 10.1. Suppose that f is a locally integrable function and μ a measure on $\mathbb{R}_+$, and that $\sigma \in \mathbb{R}$ is such that $\int_0^\infty x^{\sigma-1}|f(x)| < \infty$ and $\int_0^\infty x^{-\sigma}\,\mathrm{d}\mu < \infty$, i.e., the Mellin transforms $\widetilde{f}(s)$ and $\widetilde{\mu}(1-s)$ are defined when $\Re s = \sigma$. Suppose also that the integral $\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \widetilde{f}(s)\widetilde{\mu}(1-s)\,\mathrm{d}s$ converges at least conditionally. Suppose further that $x^\sigma f(x)$ is bounded and that f is μ-a.e. continuous. Then

$$\int_0^\infty f(x)\,\mathrm{d}\mu(x) = \frac{1}{2\pi\mathrm{i}}\int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \widetilde{f}(s)\widetilde{\mu}(1-s)\,\mathrm{d}s. \tag{10.1}$$

Although we have not found Lemma 10.1 in the literature, it is closely related to standard versions of Parseval's formula for Mellin transforms. These are usually stated for two functions, while our lemma is stated for a function and a measure; however, for absolutely continuous μ, our lemma reduces to a result for two functions. In particular, in the case that μ is absolutely continuous, Lemma 10.1 is a special case of the more general [22, Theorem 43]. (Note that we have shown in Lemma 6.1 that our measure Υ is absolutely continuous, and as discussed in Section 11.2, this extends to the measures Υ_k considered there; however, one advantage of our use of Lemma 10.1 is that we do not have to verify this by a special argument.) For completeness, we give a detailed proof below.

Note also that by a standard change of variables, as in the proof below, any version of Parseval's formula for Mellin transforms is equivalent to a corresponding version of Parseval's formula (also called Plancherel's formula) for the Fourier transform on the real line; see e.g. [22]. (From the point of view of abstract Harmonic analysis, the Mellin transform is just the Fourier transform on the multiplicative group $(\mathbb{R}_+,\cdot)$ with Haar measure $\mathrm{d}x/x$.)

Proof of Lemma 10.1. We use a well-known transformation to the Fourier transform on $\mathbb{R}$, which we do in two steps. First, let f_σ and μ_σ be the function and measure on $\mathbb{R}_+$ defined by

$$f_\sigma(x) := x^\sigma f(x), \qquad \mathrm{d}\mu_{-\sigma}(x) := x^{-\sigma}\,\mathrm{d}\mu(x) \tag{A.1}$$

(i.e., the Radon–Nikodym derivative $\mathrm{d}\mu_{-\sigma}/\,\mathrm{d}\mu = x^{-\sigma}$). Note that then $\widetilde{f_\sigma}(s) = \widetilde{f}(s+\sigma)$ and $\widetilde{\mu}_\sigma(s) = \widetilde{\mu}(s-\sigma)$, and that f_σ and μ_σ satisfy the assumptions of Lemma 10.1 with $\sigma = 0$, and so we have reduced the lemma to this case. (In particular, μ_σ is a finite measure.) Secondly, we change variables to $y \in \mathbb{R}$ by the mapping $y = \log x$; we let $F(y) := f_\sigma(e^y)$ and let ν be the (finite) measure on $\mathbb{R}$ corresponding to μ_σ. Then the assumptions on f and μ imply that $F \in L^1(\mathbb{R})$, F is bounded, F is ν-a.e. continuous, and that ν is a finite measure on $\mathbb{R}$. Furthermore, for any $\tau \in \mathbb{R}$,

(A.2)

$$\widetilde{f}(\sigma+\mathrm{i}\tau) = \int_0^\infty f(x)x^{\sigma+\mathrm{i}\tau-1}\,\mathrm{d}x = \int_0^\infty f_\sigma(x)x^{\mathrm{i}\tau-1}\,\mathrm{d}x = \int_{-\infty}^\infty F(y)e^{\mathrm{i}\tau y}\,\mathrm{d}y =: \widehat{F}(\tau),$$

the Fourier transform of F, and

(A.3)
$$\widetilde{\mu}(1-(\sigma+\mathrm{i}\tau))=\int_0^\infty x^{-\sigma-\mathrm{i}\tau}\,\mathrm{d}\mu(x)=\int_0^\infty x^{-\mathrm{i}\tau}\,\mathrm{d}\mu_\sigma(x)=\int_{-\infty}^\infty e^{-\mathrm{i}\tau y}\,\mathrm{d}\nu(y)=:\widehat{\nu}(-\tau).$$

Hence, Lemma 10.1 follows from (and is equivalent to) the following lemma for the Fourier transform on $\mathbb{R}$. $\square$

Lemma A.1. *Suppose that f is an integrable function on $\mathbb{R}$ and that ν is a finite measure on $\mathbb{R}$ such that the integral $\int_{-\infty}^\infty \widehat{f}(t)\widehat{\nu}(-t)\,\mathrm{d}t$ converges at least conditionally. Suppose further that f is bounded and ν-a.e. continuous. Then*

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}\nu(x)=\frac{1}{2\pi}\int_{-\infty}^\infty \widehat{f}(t)\widehat{\nu}(-t)\,\mathrm{d}t. \tag{A.4}$$

Proof. We begin by taking the Féjer kernel on $\mathbb{R}$, for $s>0$,

$$k_s(x):=\frac{1-\cos(sx)}{\pi s x^2}=\frac{\sin^2(sx/2)}{2\pi s(x/2)^2}, \tag{A.5}$$

which is integrable and has the Fourier transform

$$\widehat{k_s}(t)=(1-|t|/s)_+, \tag{A.6}$$

see e.g. [17, VI.(1.8), p. 124]. Thus the convolution k_s*f has Fourier transform

$$\widehat{k_s*f}(t)=\widehat{f}(t)\widehat{k_s}(t)=(1-|t|/s)_+\widehat{f}(t). \tag{A.7}$$

The function $\widehat{f}(t)$ is bounded, since $f\in L^1$, and thus (A.7) shows that $\widehat{k_s*f}$ is integrable. Consequently, the Fourier inversion formula applies and yields, for every $s>0$ and a.e. $x\in\mathbb{R}$,

$$k_s*f(x)=\frac{1}{2\pi}\int_{-\infty}^\infty e^{-\mathrm{i}tx}\widehat{f}(t)(1-|t|/s)_+\,\mathrm{d}t. \tag{A.8}$$

We obtain from (A.8) by Fubini's theorem (twice), since the double integrals are absolutely convergent by the assumptions that $f\in L^1$ and ν is finite, and thus $\widehat{f}$ and $\widehat{\nu}$ are bounded,

$$\begin{aligned}\int_{-\infty}^\infty k_s*f(x)\,\mathrm{d}\nu(x)&=\int_{-\infty}^\infty\frac{1}{2\pi}\int_{-\infty}^\infty e^{-\mathrm{i}tx}\widehat{f}(t)(1-|t|/s)_+\,\mathrm{d}t\,\mathrm{d}\nu(x)\\&=\frac{1}{2\pi}\int_{-s}^{s}\widehat{\nu}(-t)\widehat{f}(t)(1-|t|/s)\,\mathrm{d}t\\&=\frac{1}{2\pi}\int_0^1\int_{-us}^{us}\widehat{\nu}(-t)\widehat{f}(t)\,\mathrm{d}t\,\mathrm{d}u.\end{aligned} \tag{A.9}$$

The assumption that $\int_{-\infty}^\infty \widehat{f}(t)\widehat{\nu}(-t)\,\mathrm{d}t$ converges conditionally means that as $s\to\infty$, the inner integral on the last line of (A.9) tends to $J:=\int_{-\infty}^\infty \widehat{f}(t)\widehat{\nu}(-t)\,\mathrm{d}t$; it follows also that this inner integral is bounded uniformly in u and s. Hence, (A.9) and dominated convergence show that as $s \to \infty$,

$$\int_{-\infty}^{\infty} k_s * f(x) \,\mathrm{d}\nu(x) \to \frac{1}{2\pi} \int_0^1 J \,\mathrm{d}u = \frac{1}{2\pi} J. \tag{A.10}$$

Moreover, since f is bounded, it is easily seen (and well-known) that, as $s \to \infty$, $k_s * f(x) \to f(x)$ for every x such that f is continuous at x; by assumption this holds for ν-a.e. x. Since furthermore $k_s * f$ is bounded (by $\sup_x |f(x)|$), it follows by dominated convergence that

$$\int_{-\infty}^{\infty} k_s * f(x) \,\mathrm{d}\nu(x) \to \int_{-\infty}^{\infty} f(x) \,\mathrm{d}\nu(x) \qquad \text{as } s \to \infty. \tag{A.11}$$

Finally, (A.4) follows from (A.10) and (A.11). □

Remark A.2. Lemma A.1 extends, with the same proof, to complex measures ν.

Remark A.3. As far as we know, it is unknown (even in the absolutely continuous case) whether (A.4) holds for any integrable f and finite measure ν such that both integrals in (A.4) converge.

References

[1] David Aldous. Probability distributions on cladograms. In *Random Discrete Structures (Minneapolis, MN, 1993)*, volume 76 of *IMA Vol. Math. Appl.*, pages 1–18. Springer, New York, 1996.

[2] David Aldous and Svante Janson. The critical beta-splitting random tree II: Overview and open problems. arXiv 2303.02529v2, 2024.

[3] David Aldous and Svante Janson. The critical beta-splitting random tree III: The exchangeable partition representation and the fringe tree. arXiv::2412.09655, 2024.

[4] David Aldous, Svante Janson, and Xiaodan Li. The harmonic descent chain. *Electron. Commun. Probab.* 29 (2024), paper no. 77, 1–10.

[5] David Aldous and Boris Pittel. The critical beta-splitting random tree I: Heights and related results. arXiv:2302.05066, 2023. To appear in *Annals of Applied Probability.*

[6] Jean Bertoin. *Random Fragmentation and Coagulation Processes*, volume 102 of *Cambridge Studies in Advanced Mathematics.* Cambridge Univ. Press, Cambridge, 2006.

[7] Amir Dembo and Ofer Zeitouni. *Large Deviation Techniques and Applications*, 2nd ed. Springer, New York, 1998.

[8] Michael Drmota and Wojciech Szpankowski. *Analytic Information Theory. From Compression to Learning.* Cambridge Univ. Press, Cambridge, UK, 2023.

[9] Philippe Flajolet, Xavier Gourdon, and Philippe Dumas. Mellin transforms and asymptotics: harmonic sums. *Theoret. Comput. Sci.*, 144(1-2):3–58, 1995.

[10] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics.* Cambridge Univ. Press, Cambridge, UK, 2009.

[11] Alexander Gnedin and Alexander Iksanov. Regenerative compositions in the case of slow variation: a renewal theory approach. *Electron. J. Probab.*, 17:no. 77, 19 pp., 2012.

[12] Bénédicte Haas, Grégory Miermont, Jim Pitman, and Matthias Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *Ann. Probab.*, 36(5):1790–1837, 2008.

[13] Bénédicte Haas and Grégory Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron. J. Probab.* 9(4):57–97, 2004.

[14] Richard A. Handelsman and John S. Lew. Asymptotic expansion of a class of integral transforms via Mellin transforms. *Arch. Rational Mech. Anal.* **35** (1969), 382–396.

[15] Alexander Iksanov. A comment on the article 'The harmonic descent chain' by D. J. Aldous, S. Janson and X. Li. arxiv 2412.06826, 2024.
[16] Alexander Iksanov. Another proof of CLT for critical beta-splitting tree. Unpublished, 2024.
[17] Yitzhak Katznelson. *An Introduction to Harmonic Analysis*. John Wiley & Sons, New York, 1968.
[18] Brett Kolesnik. Critical beta-splitting, via contraction. arXiv 2404.16021, 2024.
[19] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST Handbook of Mathematical Functions*. U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge Univ. Press, Cambridge, 2010. Also available as *NIST Digital Library of Mathematical Functions*, `http://dlmf.nist.gov/`
[20] Ralph Neininger and Ludger Rüschendorf. On the contraction method with degenerate limit equation. *Ann. Probab.*, 32(3B):2838–2856, 2004.
[21] Richard Bruce Paris and David Kaminski. Asymptotics and Mellin–Barnes integrals. Cambridge Univ. Press, Cambridge, 2001.
[22] E. C. Titchmarsh, *Introduction to the Theory of Fourier integrals*. Oxford Univ. Press, Oxford, UK, 1937.

Department of Statistics, 367 Evans Hall # 3860, U.C. Berkeley CA 94720
Email address: `aldousdj@berkeley.edu`
URL: `www.stat.berkeley.edu/users/aldous.`

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
Email address: `svante.janson@math.uu.se`
URL: `www2.math.uu.se/∼svante`