Journal of Theoretical Probability (2026) 39:55
https://doi.org/10.1007/s10959-026-01513-5

To See the Forest for the Trees: On the Infinite Divisibility of Unlabeled Forests

Michal Bassan[1] · **Serte Donderwinkel**[2] · **Brett Kolesnik**[3]

Received: 30 July 2025 / Revised: 22 April 2026 / Accepted: 27 April 2026

Abstract

Inspired by Stufler's recent probabilistic proof of Otter's asymptotic number of unlabeled trees, we revisit work of Palmer and Schwenk, and study unlabeled forests from a probabilistic point of view. We show that the number of trees in a random forest converges, with all of its moments, to a shifted compound Poisson. We also find the asymptotic proportion of forests that are trees. The key fact is that the number of trees and the number of forests are related by a Lévy process. As such, the results by Palmer and Schwenk follow by an earlier and far-reaching limit theory by Hawkes and Jenkins. We also show how this limit theory implies results by Schwenk and by Meir and Moon, related to degrees in large random trees. Our arguments apply, more generally, to the enumeration of subexponentially weighted integer partitions, or, in fact, any setting where the underlying Lévy process follows the one big jump principle.

Keywords Asymptotic enumeration · Compound Poisson · Forest · Infinitely divisible · Integer partition · Lévy process · Lévy–Khintchine formula · Otter's constants · Pólya enumeration · Regularly varying · Subexponential · Tree · Unlabeled graph

Mathematics Subject Classification (2020) 05A16 · 05A17 · 05C05 · 05C30 · 60E07

✉ Serte Donderwinkel
s.a.donderwinkel@rug.nl

Michal Bassan
michal.bassan@keble.ox.ac.uk

Brett Kolesnik
brett.kolesnik@warwick.ac.uk

[1] Department of Statistics, University of Oxford, Oxford, United Kingdom

[2] Bernoulli Institute for Mathematics, Computer Science and AI and CogniGron (Groningen Cognitive Systems and Materials Center), University of Groningen, Groningen, The Netherlands

[3] Department of Statistics, University of Warwick, Coventry, United Kingdom

1 Introduction

Stufler [30] recently gave a probabilistic proof of Otter's [24] result on the asymptotic number t_n of unlabeled trees on n vertices, showing that

$$\lim_{n\to\infty} \alpha^n n^{5/2} t_n = \beta, \tag{1}$$

where α and β are called *Otter's constants*. We note that $\alpha \approx 0.338$ and $\beta \approx 0.534$; see, e.g., Finch [14]. The first constant satisfies

$$\alpha = \prod_{k=1}^{\infty} (1-\alpha^k)^{r_k}, \tag{2}$$

where r_n is the number of *rooted* unlabeled trees on n vertices. The second constant is related to the average preimage size of the surjection $(T, v) \mapsto R$, where T is an unlabeled tree, v is one of its n vertices, and R is T rooted at v. Specifically,

$$\lim_{n\to\infty} \frac{n t_n}{r_n} = (2\pi\beta^2)^{1/3}. \tag{3}$$

A number of years later, Palmer and Schwenk [25] studied asymptotic properties of the number f_n of unlabeled forests on n vertices, using singularity analysis. In this work, we take a probabilistic point of view. In particular, in conjunction with Stufler's [30] proof of Theorem (1), we obtain alternative, probabilistic proofs of the main results in [25], stated as (5) and (6) below. Rather than using singularity analysis, we will employ a Tauberian-type theorem by Hawkes and Jenkins [17], based on the Lévy–Khintchine formula from the theory of Lévy processes.

In fact, our methods are more widely applicable, reaching well beyond trees and forests; see Sect. 1.3 below. We will, however, focus on this specific case of interest in order to bring our probabilistic techniques and intuition into the clearest possible light.

Our first result is as follows:

Let $t(x) = \sum_{n=1}^{\infty} t_n x^n$ be the generating function for t_n, starting the sum at $n = 1$ for convenience.

We recall (see, e.g., Aldous [1]) that a random variable X is *compound Poisson* distributed, with *compounding measure* ξ, if

$$\mathbb{E}(e^{-\phi X}) = \exp\left(-\int_0^{\infty} (1 - e^{-\phi x})\xi(dx)\right), \tag{4}$$

for all $\phi > 0$.

Theorem 1 (Tree-forest distribution) *Let $\mathcal{T}_n$ be the number of trees in a uniformly random unlabeled forest on n vertices. Then, as $n \to \infty$, $\mathcal{T}_n - 1$ converges, with all of its moments, to a discrete compound Poisson random variable $\mathcal{P}$, with compounding measure $\xi_k = t(\alpha^k)/k$ on integers $k \geq 1$.*

Generalizations of this limit theorem appear in, e.g., Bell, Bender, Cameron, and Richmond [5] and Stufler [31], but not with convergence of all moments. The convergence in distribution in the theorem implies that the number of trees in the forest is tight as its size grows. We strengthen this and show that all the powers of the number of trees are uniformly integrable. Both of these facts are an effect of the well-known *one big jump principle* that implies that a forest of large size is most likely large because of one large tree, a phenomenon sometimes referred to as *condensation*.

A more general version of this result, with essentially the same proof, carries forward in any situation where t_n is replaced by some other $x_n \sim ba^{-n}n^{-c}$, with $c > 1$, or, even more generally, when $a^n x_n$ is proportional to a *subexponential* probability distribution (so that the one big jump principle applies). See Sect. 1.2 and 1.3 below for definitions and more context.

In particular, we recover Theorem 2 in [25], stating that

$$\lim_{n\to\infty} \mathbb{E}(\mathscr{T}_n) \to 1 + \sum_{k=1}^{\infty} t(\alpha^k). \tag{5}$$

In fact, by Theorem 1, and reversing the order of summation (see the end of Sect. 3.1 for details), we obtain the following result.

Corollary 2 *As* $n \to \infty$,

$$\mathbb{E}(\mathscr{T}_n) \to 1 + \sum_{k=1}^{\infty} \frac{\alpha^k}{1-\alpha^k} t_k$$

and

$$\mathrm{Var}(\mathscr{T}_n) \to \sum_{k=1}^{\infty} \frac{\alpha^k}{(1-\alpha^k)^2} t_k.$$

Using the values for t_n, $1 \le n \le 100$, available at the OEIS [29], we note that the asymptotic mean and variance of $\mathscr{T}_n$ are approximately 1.755 and 1.035, respectively.

In proving Theorem 1, we will also give an alternative, probabilistic proof of Theorem 1 in [25], concerning the asymptotic proportion of unlabeled forests that are connected, showing that

$$\lim_{n\to\infty} \frac{t_n}{f_n} \to \frac{1}{1+f(\alpha)}, \tag{6}$$

where $f(x) = \sum_{i=1}^{\infty} f_n x^n$ is the generating function for f_n; see Corollary 4 below.

1.1 Our Approach

We recall that a probability distribution π is *infinitely divisible* if, for any $n \ge 1$, there are independent and identically distributed $X_1, \ldots, X_n$ such that $X_1 + \cdots + X_n \sim \pi$ has this distribution.

The following observation that all three sequences t_n, r_n, and f_n are related by an infinitely divisible probability distribution is central to our arguments. We use the convention that $f_0 = 1$.

Theorem 3 (Otter's distribution) *Put*

$$\rho = \prod_{k=1}^{\infty} (1 - \alpha^k)^{t_k}.$$

Then, the sequence $(p_n, n \geq 0)$ *with*

$$p_n = \rho \alpha^n f_n = f_n \prod_{k=1}^{\infty} (1 - \alpha^k)^{t_k + n r_k} \tag{7}$$

is an infinitely divisible probability distribution.

We note that the proof of this theorem shows that

$$\hat{p}_n = \hat{\alpha}^n f_n \prod_{k=1}^{\infty} (1 - \hat{\alpha}^k)^{t_k} \tag{8}$$

is an infinitely divisible probability distribution for any $0 < \hat{\alpha} \leq \alpha$, so we in fact obtain a family of such distributions. The second equality in (7), however, holds only when $\hat{\alpha} = \alpha$, using (2).

Each infinitely divisible probability distribution π has an associated *Lévy process* $(L_t, t \geq 0)$, whose state $L_1 \sim \pi$ at time $t = 1$ has this distribution. The *Lévy–Khintchine formula* relates π with a corresponding *Lévy measure* ν, which controls the jumps of the process. When $(\pi_n, n \geq 0)$ is supported on the nonnegative integers, this formula states that $(\nu_j, j \geq 1)$ satisfies

$$\sum_{n=0}^{\infty} \pi_n x^n = \exp\left(- \sum_{j=1}^{\infty} (1 - x^j) \nu_j \right), \tag{9}$$

and thus, π is a compound Poisson distribution with compounding measure ν (see (4) above). In this case, the *total Lévy measure* $\lambda = \sum_{j=1}^{\infty} \nu_j$ is finite. Hence, $\mu_n = \nu_n / \lambda$ is a well-defined probability measure and π is the law of the sum of Poisson(λ) independent samples from μ.

As we will see (see Sect. 2 below), in the specific case of Otter's distribution $\pi_n = p_n$, the Lévy measure ν_n then equals

$$\tau_n = \frac{\alpha^n}{n} \sum_{d|n} d t_n, \tag{10}$$

so that by changing the order of summation (see (14) below), we find that $\rho = e^{-\lambda}$. By (1), we find that $\tau_n \sim \alpha^n t_n$ is *regularly varying* with index $\gamma = -5/2$, in the sense that $\tau_{\lfloor xn \rfloor} \sim x^{\gamma} \tau_n$, for all $x > 0$ (see, e.g., Bojanic and Seneta [6]). A general result of Hawkes and Jenkins [17], from a year before [25], states that if π_n and ν_n are related by (9) and ν_n is regularly varying with index $\gamma < -1$, then $\pi_n \sim \nu_n$. As such, we obtain the following result, which, as we will see (see again Sect. 2), implies (6).

Corollary 4 *As $n \to \infty$, we have that $f_n \sim t_n/\rho$.*

We note that p_n is also regularly varying. On the other hand, the infinitely divisible distributions $\hat{p}_n$ in (8) are not regularly varying when $\hat{\alpha} < \alpha$.

1.2 One Big Jump Principle

Our proof of Corollary 4 is based on the so-called *one big jump principle;* see, e.g., [7, 18, 23].

The result by Hawkes and Jenkins [17], as discussed above, was later extended by Embrechts and Hawkes [11], who showed that the following are equivalent, as $n \to \infty$:

(i) $\mu_n^* \sim 2\mu_n$ and $\mu_n \sim \mu_{n+1}$;
(ii) $\pi_n^* \sim 2\pi_n$ and $\pi_n \sim \pi_{n+1}$;
(iii) $\pi_n \sim \nu_n$ and $\nu_n \sim \nu_{n+1}$.

Above, as usual, μ_n^* (and, similarly, π_n^*) denotes the *self-convolution*

$$\mu_n^* = \sum_{k=0}^{n} \mu_k \mu_{n-k}.$$

A probability measure $(\mu_n, n \geq 0)$ is *subexponential* if condition (i) above is satisfied. Such distributions satisfy the so-called one big jump principle; namely, if the sum of two (or more) independent subexponential random variables takes a large value, then most likely one of them has taken (essentially) this value.

The equivalent condition (iii) then says that the probability π_n that the Lévy process $L_1 \sim \pi$ takes a large value n is asymptotically equivalent to ν_n. For $j \geq 1$, the value ν_j is the expected number of jumps by L of size j by time $t = 1$ (i.e., times $s \leq 1$ that $\lim_{\varepsilon \downarrow 0} L_{s-\varepsilon} = L_s - j$). For large j, this is close to the probability of having one jump of size j, so $\pi_n \sim \nu_n$ essentially says that if L_1 takes a large value, it is most likely due to one large jump.

Since, by (1), $\alpha^n t_n$ is regularly varying with index $\gamma = -5/2$, it is natural to expect, in light of (7) and (10), the asymptotics in Corollary 4 to hold.

In closing, let us mention that different versions of the one big jump principle have been used recently by Stufler [31] to study the sizes of the small structures in a random multiset of large total size, and by Panagioutou and Ramzews [26] to study the rare event that a multiset of large total size consists of many objects. Both of these works apply, in particular, to our current case of trees in large forests.

1.3 Further Applications

We have focused on unlabeled trees and forests; however, our methods are more widely applicable. In this way, we provide a probabilistic alternative to previous works, such as Compton [9] and Bell, Bender, Cameron, and Richmond [5], which use singularity analysis to study distributional features of random combinatorial structures.

For instance, our arguments extend immediately to the setting of *weighted integer partitions* (multisets) with subexponential weights. In fact, the potential application of this limit theory to partitions and related structures is alluded to in [17, pp. 66–67]. In the case of unlabeled trees and forests, we effectively assign weight t_k to each integer $k \geq 1$, so that there are a total of f_n weighted partitions of n. Of course, other weights are possible. See again [5, 9] and, e.g., the more recent works [2, 22, 26, 31].

In another direction, see our recent applications to the enumeration of *tournament score sequences* [4, 19], *Sinaĭ excursions* [10] and *graphical sequences* [3]. In all of these cases, in contrast to the current work, the combinatorial objects of interest are *ordered*.

Indeed, the first two are examples of renewal sequences, which can be thought of as *weighted integer compositions*. Graphical sequences, on the other hand, are *delayed* renewal sequences, with one anomalous inter-arrival time (also sometimes referred to as *Riordan arrays* in combinatorics).

We expect many further combinatorial applications to be found. The limit theory [11, 17] is based on work by Chover, Ney, and Wainger [8], which studies analytic transformations of probability measures, vastly generalizing the classical renewal theorem. As such, our current methods can, subject to the conditions in [11, 17], be applied when a combinatorial structure decomposes into smaller irreducible parts, as is often the case.

Finally, let us highlight the vast generalization of the main limit theorems in [11, 17], stated as Theorem C in Embrechts and Omey [12], where $\exp(z)$ in (9) is replaced with some other analytic function $\phi(z)$, pointing to applications awaiting, well beyond the Lévy–Khintchine correspondence.

1.4 Outline

In Sect. 2, we prove Theorem 3 and its Corollary 4. In Sect. 3, we prove our main result, Theorem 1. Finally, in Sect. 4, we discuss some further applications, related to cycle index asymptotics and degrees in large random trees, giving an alternative proof of a result of Schwenk [28] (cf. Meir and Moon [21]).

1.5 Acknowledgments

MB is supported by a Clarendon Fund Scholarship. SD acknowledges the financial support of the CogniGron research center and the Ubbo Emmius Funds (University of Groningen). Her research was also partially supported by the Marie Skłodowska-Curie grant GraPhTra (Universality in phase transitions in random graphs), grant agreement ID 101211705. Part of this research was performed, while MB and SD were visiting the Mathematical Sciences Research Institute (MSRI), now becoming the Simons Laufer Mathematical Sciences Institute (SLMath), which is supported by the National Science Foundation (Grant No. DMS-1928930).

2 Otter's Distribution

A forest is a multiset of trees. As such, their generating functions $t(x)$ and $f(x)$ are related via *Pólya's exponential,* as can be seen using the *log-exp transform,* as a simple instance of the *symbolic method* in Flajolet and Sedgewick [15, p. 29]. Indeed, as explained therein, let $\mathscr{T}$ be the set of all unlabeled trees. For each $T \in \mathscr{T}$, let $|T| \geq 1$ be its number of vertices. Each forest can be expressed as a sequence of trees, in some canonical order. As such, it follows that $1 + f(x)$ (with 1 accounting for the $f_0 = 1$ empty forest containing no trees) is an infinite product of geometric series, with one for each $T \in \mathscr{T}$:

$$\begin{aligned} 1 + f(x) &= \prod_{T \in \mathscr{T}} \frac{1}{1 - x^{|T|}} = \prod_{k=1}^{\infty} \left(\frac{1}{1 - x^k} \right)^{t_k} \\ &= \exp\left(-\sum_{k=1}^{\infty} t_k \log(1 - x^k) \right) = \exp\left(\sum_{k=1}^{\infty} \frac{t(x^k)}{k} \right). \end{aligned} \tag{11}$$

Technically, since $\mathscr{T}$ is infinite, a limiting procedure is required in order to make the above precise, but this is the basic idea.

The coefficient of x^n inside the exponential, on the right-hand side of (11), is given by

$$[x^n] \sum_{k=1}^{\infty} \frac{t(x^k)}{k} = \frac{1}{n} \sum_{d|n} d t_d. \tag{12}$$

Therefore, the Lévy–Khintchine formula (9) is satisfied, with infinitely divisible probability distribution

$$p_n = \alpha^n f_n e^{-\lambda}$$

and Lévy measure

$$\tau_n = \frac{\alpha^n}{n} \sum_{d|n} d t_d \sim \alpha^n t_n, \tag{13}$$

which, upon interchanging the order of summation, can be seen to have total measure

$$\lambda = \sum_{n=1}^{\infty} \frac{\alpha^n}{n} \sum_{d|n} d t_d = \sum_{k=1}^{\infty} t_k \sum_{i=1}^{\infty} \frac{\alpha^{ki}}{i} = -\sum_{k=1}^{\infty} t_k \log(1 - \alpha^k). \tag{14}$$

Noting that

$$e^{-\lambda} = \prod_{k=1}^{\infty} (1 - \alpha^k)^{t_k} = \rho,$$

Theorem 3 follows.

Finally, as discussed above, by applying [17] we find that $p_n \sim \tau_n$ and Corollary 4 follows. In light of (11), this is equivalent to (6).

3 Tree-Forest Distribution

We recall that the *cycle index* $Z(S_n)$ of the symmetric group S_n is the polynomial in the variables $x_1, \dots, x_n$ defined by

$$Z(S_n) = Z(S_n; x_1, \dots, x_n) = \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{\ell=1}^{n} x_\ell^{c_\ell(\sigma)},$$

where $c_\ell(\sigma)$ is the number of cycles in the permutation $\sigma \in S_n$ of length ℓ. It can be computed recursively. We note that $Z_0 = 1$ and, for $k \geq 1$,

$$kZ(S_k) = \sum_{i=1}^{k} x_i Z(S_{k-i}), \tag{15}$$

which can be deduced by considering the cycle of 1 in a $\sigma \in S_k$. By classical Pólya theory (see, e.g., [15, Eq. (91), p. 86]) for any $k \geq 1$, the generating function for the number of unlabelled forests with k trees equals $Z(S_k, t(x), \dots, t(x^k))$, so the number of unlabelled forests of size n with exactly $k+1$ trees equals

$$[x^n]Z(S_{k+1}, t(x), \dots, t(x^{k+1})).$$

By Robinson and Schwenk [27], this is asymptotically equivalent to $t_n \zeta_k$ as $n \to \infty$, with $\zeta_0 = 1$ and

$$\zeta_k = Z(S_k, t(\alpha), \dots, t(\alpha^k)), \quad k \geq 1. \tag{16}$$

As observed in [25, p. 120], dividing by f_n and using its asymptotics in Corollary 4, the probability that a random unlabeled forest on n vertices has exactly $k+1$ trees converges to

$$\gamma_k := \rho \zeta_k.$$

as $n \to \infty$.

Moreover, by (15), the generating functions $\zeta(x) = \sum_{k=0}^{\infty} \zeta_k x^k$ and $\xi(x) = \sum_{i=1}^{\infty} t(\alpha^i) x^i$ satisfy $x\zeta'(x) = \xi(x)\zeta(x)$. By integrating, we find that

$$\sum_{k=0}^{\infty} \zeta_k x^k = \exp\left(\sum_{i=1}^{\infty} \frac{t(\alpha^i)}{i} x^i \right). \tag{17}$$

Comparing (17) with (9) and (11), it follows that the sequence $(\gamma_k, k \geq 0)$ is an infinitely divisible probability distribution, with Lévy measure $(\nu_i, i \geq 1)$ given by $\nu_i = t(\alpha^i)/i$. Therefore, we obtain the following result, by the discrete compound Poisson characterization of infinitely divisible probability distributions on the non-negative integers; see, e.g., Feller [13, p. 290].

Proposition 5 *Let $\mathscr{T}_n$ be the number of trees in a uniformly random unlabeled forest on n vertices. Then, as $n \to \infty$, $\mathscr{T}_n - 1$ converges to a discrete compound Poisson random variable $\mathscr{P}$, with compounding measure $t(\alpha^k)/k$ on integers $k \geq 1$.*

In the next section, we show that this probabilistic rephrasing of the observations in [25] yields a direct proof of (5) that appears as Theorem 2 in [25], as well as an extension to the higher moments of $\mathscr{T}_n$.

3.1 Convergence of Moments

With Proposition 5 in hand, we finish the proof of Theorem 1 by showing that $\mathscr{T}_n$ and its powers are uniformly integrable.

The proof is once again related to the one big jump principle: A forest with many trees can be partitioned into multiple large forests, which violates that a forest is in fact most likely big because it has one big tree.

First, we provide the details of the uniform integrability of $\mathscr{T}_n$ and will then explain how to adapt the argument to higher powers.

Lemma 6 *We have that*

$$\lim_{K\to\infty} \limsup_{n\to\infty} \mathbb{E}[\mathscr{T}_n \mathbf{1}_{\{\mathscr{T}_n > K\}}] = 0.$$

Proof Fix K and observe that

$$\mathbb{E}[\mathscr{T}_n \mathbf{1}_{\{\mathscr{T}_n > K\}}] = K\mathbb{P}(\mathscr{T}_n > K) + \sum_{k=K}^{\infty} \mathbb{P}(\mathscr{T}_n > k). \tag{18}$$

Since $\mathscr{T}_n$ converges in distribution to $\mathscr{P} + 1$, it follows that

$$K\mathbb{P}(\mathscr{T}_n > K) \to K\mathbb{P}(\mathscr{P} \geq K).$$

Since $\mathbb{E}(\mathscr{P}) < \infty$, the right-hand side can be made arbitrarily small by making K large.

We turn to the second term in (18). First, we note, by (1) and Corollary 4, that there are constants $c, C > 0$ such that

$$c \leq \alpha^n n^{5/2} f_n \leq C, \tag{19}$$

for all $n \geq 1$. Next, we observe that, for any n and k, and any forest with n vertices and more than k trees, we can decompose the forest into a pair of forests, such that: (1) their sizes sum to n, (2) the first has at least $\lfloor k/2 \rfloor$ vertices, and (3) the second is at least as big as the first. As such, the number of forests with n vertices and more than k trees is at most

$$\sum_{\ell=\lfloor k/2 \rfloor}^{\lfloor n/2 \rfloor} f_\ell f_{n-\ell}.$$

Therefore, applying (19), we find that $\mathbb{P}(\mathscr{T}_n > k)$ is at most

$$\frac{C^2}{c}\sum_{\ell=\lfloor k/2\rfloor}^{\lfloor n/2\rfloor}\left(\frac{n}{\ell(n-\ell)}\right)^{5/2} \le \frac{2^{5/2}C^2}{c}\sum_{\ell=\lfloor k/2\rfloor}^{\lceil n/2\rceil}\frac{1}{\ell^{5/2}} \le \frac{c'}{k^{3/2}},$$

for some constant c' not depending on k and n. This is summable, so, for any $\varepsilon > 0$, we can choose K large so that $\mathbb{E}[\mathscr{T}_n\mathbf{1}_{\{\mathscr{T}_n>K\}}] < \varepsilon$, for all large n. □

The proof above can be adapted to show that $\mathscr{T}_n^m$ is uniformly integrable, for any integer $m \ge 1$. The basic idea is to split a forest with more than k trees into $m+1$ forests of size at least $\lfloor k/(m+1)\rfloor$ and repeat the arguments above. The adaptation is straightforward, so we omit the details. Together with Proposition 5, this yields Theorem 1.

The argument also carries over to regularly varying distributions π_n with index $\gamma < -1$ by making enough splits to make the resulting power series summable (this covers, e.g., the case of forests of rooted trees, as then $\gamma = -3/2$). More generally, for a subexponential distribution π_n that is bounded from above by a regularly varying distribution with index $\gamma < -1$, we may use [31, Lemma 2.6.2] (see also [16, Theorems 4.8 and 4.30]) with $g(z) = \sum_{n\ge\lfloor \frac{k}{m_1}\rfloor}\pi_n z^n$ and $f(z) = z^{m+1}$ to bound the number of sequences of $m+1$ elements of size at least $\lfloor k/(m+1)\rfloor$ and then proceed as in the regularly varying case.

Proof Corollary 2 To see that Theorem 1 implies Corollary 2, we first note (see, e.g., Mane [20, p. 279]) that

$$\mathbb{E}(\mathscr{P}) = \sum_{k=1}^{\infty} t(\alpha^k).$$

and

$$\mathrm{Var}(\mathscr{P}) = \sum_{k=1}^{\infty} k t(\alpha^k).$$

By similar reasoning as (12), we see that

$$[x^n]\sum_{k=1}^{\infty} k^p t(x^k) = \sum_{d|n}(n/d)^p t_d.$$

Reversing the order of summation, as in (14) above, we find that

$$\mathbb{E}(\mathscr{P}) = \sum_{k=1}^{\infty}\alpha^k\sum_{d|k} t_d = \sum_{k=1}^{\infty} t_k\sum_{i=1}^{\infty}\alpha^{ki} = \sum_{k=1}^{\infty}\frac{\alpha^k}{1-\alpha^k}t_k$$

and

$$\mathrm{Var}(\mathscr{P}) = \sum_{k=1}^{\infty} k\alpha^k\sum_{d|k}\frac{t_d}{d} = \sum_{k=1}^{\infty} t_k\sum_{i=1}^{\infty} i\alpha^{ki} = \sum_{k=1}^{\infty}\frac{\alpha^k}{(1-\alpha^k)^2}t_k,$$

as claimed. □

4 Cycle Index Asymptotics

We conclude with two related applications of [11, 17] to the asymptotics of cycle indices of the symmetric group.

4.1 Degrees in Large Random Trees

Recall the cycle index ζ_n, as defined in (16). Note that $\alpha^{-n}t(\alpha^n) \to t_1 = 1$, as $n \to \infty$. Therefore, by (17) and Theorem 4.2 in [17] (where, for this application, we set $a_n = \alpha^{-n}t(\alpha^n)$ and $b_n = \alpha^{-n}\zeta_n$), it follows that, as $n \to \infty$,

$$\zeta_n \sim e^{\xi}\alpha^n, \tag{20}$$

where

$$\xi = \sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{t(\alpha^k)}{\alpha^k} - 1\right).$$

Noting again that $t_1 = 1$, and reversing the order of summation, we see that

$$e^{\xi} = \prod_{k=2}^{\infty}(1 - \alpha^{k-1})^{-t_k}.$$

This result was proved by Schwenk [28], using an argument that involves a certain generalization of the inclusion–exclusion principle. Another proof was later given by Meir and Moon [21]. In fact, as noted above, the result is a direct consequence of [17] that precedes the latter.

The asymptotics (20) are used, together with [27], to obtain information about degrees in large random trees; see Corollary 4.1 in [28].

4.2 Subexponential Sequences

Finally, by combining [11, Theorem 1] and (15), we obtain the following generalization of Corollary 4.

Theorem 7 *Suppose that* $(\nu_k, k \geq 1)$ *is a positive sequence, with a finite total sum* $\lambda = \sum_{k=1}^{\infty} \nu_k < \infty$. *Suppose also that* $\mu_k = \nu_k/\lambda$ *defines a subexponential probability distribution on the integers* $k \geq 1$. *Then, as* $n \to \infty$,

$$Z(S_n; \nu_1, 2\nu_2, \ldots, n\nu_n) \sim e^{\lambda}\nu_n. \tag{21}$$

In this work, we have demonstrated a procedure for obtaining the asymptotics of a sequence $(1 = a_0, a_1, \ldots)$. Suppose, for some $\gamma > 0$, that $\gamma^n a_n$ is summable, so proportional to a probability distribution π_n. Then, if (9) holds for some positive sequence $(\nu_n, n \geq 1)$, with total sum λ, and $\mu_n = \nu_n/\lambda$ is subexponential, then $\pi_n \sim \nu_n$. The above theorem reverses this procedure by, given ν_n, identifying π_n that satisfies (9). Indeed, for $\pi_n = e^{-\lambda} Z_n$ and Z_n, the cycle index in (21) above, (9) follows by setting $x_i = i\nu_i$ in (15) above. Thus, if $\mu_n = \nu_n/\lambda$ is subexponential, then $\pi_n \sim \nu_n$ and the theorem follows.

Author Contributions Both the research and the writing of the manuscript were conducted collaboratively with all three authors.

Data Availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Aldous, D.: Probability Approximations Via the Poisson Clumping Heuristic, Applied Mathematical Sciences. Springer-Verlag, New York (1989)
2. Barbour, A.D., Granovsky, B.L.: Random combinatorial structures: the convergent case. J. Combin. Theory Ser. A **109**(2), 203–220 (2005)
3. Bassan, M., Donderwinkel, S., Kolesnik, B.: Graphical sequences and plane trees, Comb. Probab. Comput. (2026). published online https://doi.org/10.1017/S0963548325100345
4. Bassan, M., Donderwinkel, S., Kolesnik, B.: Tournament score sequences, Erdős-Ginzburg-Ziv numbers, and the Lévy-Khintchine method. Electron. Commun. Probab. **31**, 3 (2026)
5. Bell, J.A., Bender, E.A., Cameron, P.J., Richmond, L.B.: Asymptotics for the probability of connectedness and the distribution of number of components. Electron. J. Combin. **7**(R33), 1–22 (2000)
6. Bojanic, R., Seneta, E.: A unified theory of regularly varying sequences. Math. Z. **134**, 91–106 (1973)
7. Chistyakov, V.P.: A theorem on sums of independent random variables and its applications to branching random processes. Theory Probab. Appl. **9**(4), 640–648 (1964)
8. Chover, J., Ney, P., Wainger, S.: Functions of probability measures. J. Analyse Math. **26**, 255–302 (1973)
9. Compton, K.J.: Some methods for computing component distribution probabilities in relational structures. Discrete Math. **66**(1), 59–77 (1987)
10. Donderwinkel, S., Kolesnik, B.: Sinaĭ excursions: an analogue of Sparre Andersen's formula for the area process of a random walk. J. Stat. Phys. **193**(2), 11 (2026)
11. Embrechts, P., Hawkes, J.: A limit theorem for the tails of discrete infinitely divisible laws with applications to fluctuation theory. J. Austral. Math. Soc. Ser. A **32**(3), 412–422 (1982)
12. Embrechts, P., Omey, E.: Functions of power series. Yokohama Math. J. **32**(1–2), 77–88 (1984)
13. Feller, W.: An introduction to probability theory and its applications, vol. I, 3rd edn. John Wiley & Sons, Inc., New York-London-Sydney (1968)
14. Finch, S.R.: Mathematical constants, Encyclopedia Math. Appl., vol. 94. Cambridge University Press, Cambridge (2003)
15. Flajolet, P., Sedgewick, R.: Analytic combinatorics. Cambridge University Press, Cambridge (2009)
16. Foss, S., Korshunov, D., Zachary, S.: An Introduction to heavy-tailed and subexponential distributions, 2nd edn. Springer, New York (2013)
17. Hawkes, J., Jenkins, J.D.: Infinitely divisible sequences. Scand. Actuar. J. **1978**(2), 65–76 (1978)

18. Heyde, C.: On large deviation probabilities in the case of attraction to a non-normal stable law. Sankhyā A, 253–258 (1968)
19. Kolesnik, B.: The asymptotic number of score sequences. Combinatorica **43**(4), 827–844 (2023)
20. Mane, S.R.: Compound poisson distributions: a general framework. J. Indian Soc. Probab. Stat. **26**(1), 277–338 (2025)
21. Meir, A., Moon, J.W.: On an asymptotic evaluation of the cycle index of the symmetric group. Discrete Math. **46**(1), 103–105 (1983)
22. Mutafchiev, L.: Limit theorems for the number of parts in a random weighted partition. Electron. J. Combin. **18**(1), #P206, 1–27 (2011)
23. Nagaev, A.V.: Integral limit theorems taking large deviations into account when Cramér's condition does not hold. Theory Probab. Appl. **14**(1), 51–64 (1969)
24. Otter, R.: The Number of Trees. Ann. of Math. (2) **49**, 583–599 (1948)
25. Palmer, E.M., Schwenk, A.J.: On the number of trees in a random forest. J. Combin. Theory Ser. B **27**(2), 109–121 (1979)
26. Panagiotou, K., Ramzews, L.: Asymptotic enumeration and limit laws for multisets: the subexponential case. Ann. Inst. Henri Poincaré Probab. Stat. **60**(1), 612–635 (2024)
27. Robinson, R.W., Schwenk, A.J.: The distribution of degrees in a large random tree. Discrete Math. **12**(4), 359–372 (1975)
28. Schwenk, A.J.: An asymptotic evaluation of the cycle index of a symmetric group. Discrete Math. **18**(1), 71–78 (1977)
29. Sloane, N.J.A.: Number of trees with n unlabeled nodes. The on-line encyclopedia of integer sequences, entry https://oeis.org/A000055
30. Stufler, B.: Probabilistic enumeration and equivalence of nonisomorphic trees. Discrete Math. Theor. Comput. Sci. **27**(3), 18 (2025)
31. Stufler, B.: Unlabelled Gibbs partitions. Combin. Probab. Comput. **29**(2), 293–309 (2020)