

On moment sequences and mixed Poisson distributions

Markus Kuba

*Institute of Applied Mathematics and Natural Sciences
University of Applied Sciences - Technikum Wien
Höchstädtplatz 5, 1200 Wien
e-mail: kuba@technikum-wien.at*

and

Alois Panholzer*

*Institut für Diskrete Mathematik und Geometrie
Technische Universität Wien
Wiedner Hauptstr. 8-10/104
1040 Wien, Austria
e-mail: Alois.Panholzer@tuwien.ac.at*

Abstract: In this article we survey properties of mixed Poisson distributions and probabilistic aspects of the Stirling transform: given a non-negative random variable X with moment sequence $(\mu_s)_{s \in \mathbb{N}}$ we determine a discrete random variable Y , whose moment sequence is given by the Stirling transform of the sequence $(\mu_s)_{s \in \mathbb{N}}$, and identify the distribution as a mixed Poisson distribution. We discuss properties of this family of distributions and present a new simple limit theorem based on expansions of factorial moments instead of power moments. Moreover, we present several examples of mixed Poisson distributions in the analysis of random discrete structures, unifying and extending earlier results. We also add several entirely new results: we analyse triangular urn models, where the initial configuration or the dimension of the urn is not fixed, but may depend on the discrete time n . We discuss the branching structure of plane recursive trees and its relation to table sizes in the Chinese restaurant process. Furthermore, we discuss root isolation procedures in Cayley trees, a parameter in parking functions, zero contacts in lattice paths consisting of bridges, and a parameter related to cyclic points and trees in graphs of random mappings, all leading to mixed Poisson-Rayleigh distributions. Finally, we indicate how mixed Poisson distributions naturally arise in the critical composition scheme of Analytic Combinatorics.

MSC 2010 subject classifications: 60C05.

Keywords and phrases: Mixed Poisson distribution, factorial moments, Stirling transform, limiting distributions, urn models, parking functions, record-subtrees.

Received September 2014.

*AP was supported by the Austrian Science Foundation FWF, grant P25337-N23.

Contents

1	Introduction	90
1.1	Motivation	91
1.2	Notation and terminology	92
1.3	Plan of the paper	92
2	Moment sequences and mixed Poisson distributions	93
2.1	Discrete distributions and factorial moments	93
2.2	Properties of mixed Poisson distributions	93
2.3	The method of moments and basic limit laws	98
3	Examples and applications	103
3.1	Block sizes in k -Stirling permutations	103
3.2	Diminishing Pólya-Eggenberger urn models	106
3.3	Descendants in increasing trees.	109
3.4	Node-degrees in plane recursive trees.	113
3.5	Branching structures in plane recursive trees	115
3.6	Distribution of table sizes in the Chinese restaurant process	117
4	Triangular urn models	121
5	Mixed Poisson-Rayleigh laws	125
5.1	The number of inversions in labelled tree families	125
5.2	Record-subtrees in Cayley trees	127
5.3	Edge-cutting in Cayley trees	132
5.4	Parking functions and growth of the initial cluster	138
5.5	Zero contacts in bridges	143
5.6	Cyclic points and trees in graphs of random mappings	146
6	Multivariate mixed Poisson distributions	148
7	Outlook and extensions	149
7.1	Mixed Poisson distributions in Analytic Combinatorics - compositions	149
7.2	Extensions and open problems	150
	Acknowledgements	151
	References	151

1. Introduction

In combinatorics the *Stirling transform* of a given sequence $(a_s)_{s \in \mathbb{N}}$, see [12, 75], is the sequence $(b_s)_{s \in \mathbb{N}}$, with elements given by

$$b_s = \sum_{k=1}^s \left\{ \begin{matrix} s \\ k \end{matrix} \right\} a_k, \quad s \geq 1. \quad (1.1)$$

The inverse Stirling transform of the sequence $(b_s)_{s \in \mathbb{N}}$ is obtained as follows:

$$a_s = \sum_{k=1}^s (-1)^{s-k} \left[\begin{matrix} s \\ k \end{matrix} \right] b_k, \quad s \geq 1. \quad (1.2)$$

Here $\left\{ \begin{smallmatrix} s \\ k \end{smallmatrix} \right\}$ denote the Stirling numbers of the second kind, counting the number of ways to partition a set of s objects into k non-empty subsets, see [73] or [30], and $\left[\begin{smallmatrix} s \\ k \end{smallmatrix} \right]$ denotes the unsigned Stirling numbers of the first kind, counting the number of permutations of s elements with k cycles [30]. These numbers appear as coefficients in the expansions

$$x^s = \sum_{k=0}^s \left\{ \begin{smallmatrix} s \\ k \end{smallmatrix} \right\} x^k, \quad x^{\underline{s}} = \sum_{k=0}^s (-1)^{s-k} \left[\begin{smallmatrix} s \\ k \end{smallmatrix} \right] x^k, \quad (1.3)$$

relating ordinary powers x^s to the so-called falling factorials $x^{\underline{s}} = x(x-1)\dots(x-(s-1))$, $s \in \mathbb{N}_0$. On the level of exponential generating functions $A(z) = \sum_{s \geq 1} a_s z^s / s!$ and $B(z) = \sum_{s \geq 1} b_s z^s / s!$, the Stirling transform and the relations (1.1) and (1.2) turn into

$$B(z) = A(e^z - 1), \quad A(z) = B(\log(1 + z)). \quad (1.4)$$

This definition is readily generalized: given a sequence $(a_s)_{s \in \mathbb{N}}$ the generalized Stirling transform with parameter $\rho > 0$ is the sequence $(b_s)_{s \in \mathbb{N}}$ with

$$b_s = \sum_{k=1}^s \rho^k \left\{ \begin{smallmatrix} s \\ k \end{smallmatrix} \right\} a_k, \quad \text{such that } a_s = \frac{1}{\rho^s} \sum_{k=1}^s (-1)^{s-k} \left[\begin{smallmatrix} s \\ k \end{smallmatrix} \right] b_k, \quad s \geq 1. \quad (1.5)$$

On the level of exponential generating functions: $B(z) = A(\rho(e^z - 1))$ and $A(z) = B(\log(1 + \frac{z}{\rho}))$. The aim of this work is to discuss several probabilistic aspects of a generalized Stirling transform with parameter $\rho > 0$ in connection with moment sequences and *mixed Poisson distributions*, pointing out applications in the analysis of random discrete structures. Given a non-negative random variable X with power moments $\mathbb{E}(X^s) = \mu_s \in \mathbb{R}^+$, $s \geq 1$, we study the properties of another random variable Y , given its sequence of *factorial moments* $\mathbb{E}(Y^{\underline{s}}) = \mathbb{E}(Y(Y-1)\dots(Y-(s-1)))$, which are determined by the moments of X ,

$$\mathbb{E}(Y^{\underline{s}}) = \rho^s \mathbb{E}(X^s) = \rho^s \mu_s, \quad s \geq 1, \quad (1.6)$$

where $\rho > 0$ denotes an auxiliary scale parameter. Moreover, we discuss relations between the moment generating functions $\psi(z) = \mathbb{E}(e^{zX})$ and $\varphi(z) = \mathbb{E}(e^{zY})$ of X and Y , respectively.

1.1. Motivation

Our main motivation to study random variables with a given sequence of factorial moments (1.6) stems from the analysis of combinatorial structures. In many cases, amongst others the analysis of inversions in labelled tree families [64], stopping times in urn models [51, 53], node degrees in increasing trees [49], block sizes in k -Stirling permutations [51], descendants in increasing trees [47], ancestors and descendants in evolving k -tree models [65], pairs of random variables X and Y arise as limiting distributions for certain parameters of interest associated

to the combinatorial structures. The random variable X can usually be determined via its (power) moment sequence $(\mu_s)_{s \in \mathbb{N}}$, and the random variable Y in terms of the sequence of factorial moments satisfying relation (1.6). An open problem was to understand in more detail the nature of the random variable Y . In [53, 64] a few results in this direction were obtained. The goal of this work is twofold: first, to survey the properties of mixed Poisson distributions, and second to discuss their appearances in combinatorics and the analysis of random discrete structures, complementing existing results; in fact, we will add several entirely new results. It will turn out that the identification of the distribution of Y can be directly solved using mixed Poisson distributions, which are widely used in applied probability theory, see for example [20, 42, 59, 60, 77]. In the analysis of random discrete structures, mixed Poisson distributions have been used mainly in the context of Poisson approximation, see e.g. [31]. In this work we point out the appearance of mixed Poisson distributions as a genuine limiting distribution, and also present closely related phase transitions. In particular, we discuss natural occurrences of mixed Poisson distributions in urn models of a *non-standard nature* – either the size of the urn, or the initial conditions are allowed to depend on the discrete time.

1.2. Notation and terminology

We denote with \mathbb{R}^+ the non-negative real numbers. Here and throughout this work we use the notation $x^{\underline{s}} = x(x-1) \dots (x-s+1)$ for the falling factorials, and $x^{\overline{s}} = x(x+1) \dots (x+s-1)$ for the rising factorials.¹ Moreover, we denote with $\left\{ \begin{smallmatrix} s \\ k \end{smallmatrix} \right\}$ the Stirling numbers of the second kind. We use the notation $U \stackrel{\mathcal{L}}{=} V$ for the equality in distribution of random variables U and V , and $U_n \xrightarrow{\mathcal{L}} V$ denotes the convergence in distribution of a sequence of random variables U_n to V . The indicator variable of the event A is denoted by $\mathbf{1}_A$. Throughout this work the term “convergence of all moments” of a sequence of random variables refers exclusively to the convergence of all non-negative integer moments. Given a formal power series $f(z) = \sum_{n \geq 0} a_n z^n$ we denote with $[z^n]$ the extraction of coefficients operator: $[z^n]f(z) = a_n$. Furthermore, we denote with E_v the evaluation operator of the variable v at the value $v = 1$, and with D_v the differential operator with respect to v .

1.3. Plan of the paper

In the next section we state the definition of mixed Poisson distributions and discuss its properties. In Section 3 we collect several examples from the literature, unifying and extending earlier results. Furthermore, in Section 4 we present a novel approach to balanced triangular urn models and its relation to mixed Poisson distributions. Section 5 is devoted to new results concerning mixed Poisson

¹The notation $x^{\underline{s}}$ and $x^{\overline{s}}$ was introduced and popularized by Knuth; alternative notations for the falling factorials include the Pochhammer symbol $(x)_s$, which is unfortunately sometimes also used for the rising factorials.

distributions with Rayleigh mixing distribution; in particular, we discuss node isolation in Cayley trees, zero contacts in directed lattice paths, and also cyclic points in random mappings. Finally, in Section 6 we discuss multivariate mixed Poisson distributions.

2. Moment sequences and mixed Poisson distributions

2.1. Discrete distributions and factorial moments

In order to obtain a random variable Y with a prescribed sequence of factorial moments, given according to Equation (1.6) by $\mathbb{E}(Y^{\underline{s}}) = \rho^s \mu_s$, a first *ansatz* would be the following. Let Y denote a discrete random variable with support the non-negative integers, and $p(v)$ its probability generating function,

$$p(v) = \mathbb{E}(v^Y) = \sum_{\ell \geq 0} \mathbb{P}\{Y = \ell\} v^\ell.$$

The factorial moments of Y can be obtained from the probability generating function by repeated differentiation,

$$\mathbb{E}(Y^{\underline{s}}) = \sum_{\ell \geq 0} \ell^{\underline{s}} \mathbb{P}\{Y = \ell\} = E_v D_v^s p(v), \quad s \geq 0. \quad (2.1)$$

Consequently, we can describe the probability mass function of the random variable Y as follows:

$$p(v) = \sum_{s \geq 0} \mathbb{E}(Y^{\underline{s}}) \frac{(v-1)^s}{s!} = \sum_{s \geq 0} (v-1)^s \frac{\mu_s \rho^s}{s!} = \sum_{\ell \geq 0} v^\ell \sum_{s \geq \ell} \binom{s}{\ell} (-1)^{s-\ell} \frac{\mu_s \rho^s}{s!}.$$

This implies that

$$\mathbb{P}\{Y = \ell\} = [v^\ell] p(v) = \sum_{s \geq \ell} \binom{s}{\ell} (-1)^{s-\ell} \frac{\mu_s \rho^s}{s!}, \quad \ell \geq 0. \quad (2.2)$$

Up to now the calculations have been purely symbolic, no convergence issues have been addressed. In order to put the calculations above on solid grounds, and to identify the distribution, we discuss mixed Poisson distributions and their properties in the next subsection.

2.2. Properties of mixed Poisson distributions

Definition 1. Let X denote a non-negative random variable, with cumulative distribution function $\Lambda(\cdot)$, then the discrete random variable Y with probability mass function given by

$$\mathbb{P}\{Y = \ell\} = \frac{1}{\ell!} \int_{\mathbb{R}_+} X^\ell e^{-X} d\Lambda, \quad \ell \geq 0,$$

has a mixed Poisson distribution with mixing distribution X , in symbol $Y \stackrel{\mathcal{L}}{=} \text{MPo}(X)$.

The boundary case $X \stackrel{\mathcal{L}}{=} 0$ leads to a degenerate distribution with all mass concentrated at zero. A more compact notation for the probability mass function of Y is sometimes used instead of the one given above, namely $\mathbb{P}\{Y = \ell\} = \frac{1}{\ell!} \mathbb{E}(X^\ell e^{-X})$. One often encounters a slightly different definition, which includes a scale parameter $\rho \geq 0$:

$$\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} \int_{\mathbb{R}_+} X^\ell e^{-\rho X} d\Lambda, \quad \ell \geq 0,$$

or $\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} \mathbb{E}(X^\ell e^{-\rho X})$. This corresponds to a scaling of the mixing distribution, $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$. Here and throughout this work we call $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ a mixed Poisson distributed random variable with mixing distribution X and scale parameter ρ .

Example 1. The ordinary Poisson distribution $Y \stackrel{\mathcal{L}}{=} \text{Po}(\rho)$ with parameter $\rho > 0$,

$$\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} e^{-\rho}, \quad \ell \geq 0,$$

arises as a mixed Poisson distribution with degenerate mixing distribution $X \stackrel{\mathcal{L}}{=} 1$.

Example 2. The negative binomial distribution $Y \stackrel{\mathcal{L}}{=} \text{NegBin}(r, p)$ with parameters $p \in [0, 1)$ and $r > 0$,

$$\mathbb{P}\{Y = \ell\} = \binom{\ell + r - 1}{\ell} p^\ell (1 - p)^r, \quad \ell \geq 0,$$

arises as a mixed Poisson distribution with a Gamma mixing distribution $X \stackrel{\mathcal{L}}{=} \text{Gamma}(r, \theta)$ scaled by $\rho \geq 0$, such that the parameters θ and ρ satisfy $\theta \cdot \rho = p/(1 - p)$. In particular, for $\theta = 1$ the parameter p is given by $p = \rho/(1 + \rho)$. A special instance of this class of distributions is the geometric distribution $\text{Geom}(p) = \text{NegBin}(1, p)$.

Note that a Gamma distributed r.v. $X \stackrel{\mathcal{L}}{=} \text{Gamma}(r, \theta)$ has the probability density function

$$f(x; r, \theta) = \frac{x^{r-1} e^{-\frac{x}{\theta}}}{\theta^r \Gamma(r)}, \quad x > 0.$$

Example 3. A Rayleigh distributed r.v. $X \stackrel{\mathcal{L}}{=} \text{Rayleigh}(\sigma)$ with parameter σ has the probability density function

$$f(x; \sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0,$$

and is fully characterized by its (power) moment sequence:

$$\mathbb{E}(X^s) = \sigma^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right).$$

A discrete random variable Y with probability mass function

$$\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} \int_0^\infty x^{\ell+1} e^{-\rho x - \frac{x^2}{2}} dx, \quad \ell \geq 0,$$

arises as a mixed Poisson distribution $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ with mixing distribution $X \stackrel{\mathcal{L}}{=} \text{Rayleigh}(1)$ and scale parameter ρ . We call Y a Poisson-Rayleigh distribution with parameter ρ . Note that for $\rho < 1$ we can expand $e^{-\rho x}$ and obtain a series representation of $\mathbb{P}\{Y = \ell\}$. Another representation valid for all $\rho > 0$ can be stated in terms of the incomplete gamma function $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$:

$$\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} e^{\frac{\rho^2}{2}} \sum_{i=0}^{\ell+1} \binom{\ell+1}{i} (-\rho)^{\ell+1-i} 2^{\frac{i-1}{2}} \Gamma\left(\frac{i+1}{2}, \frac{\rho^2}{2}\right).$$

Example 4. The Neyman Type A Distribution is a discrete probability distribution often used in biology and ecology [59, 60]. It is a mixed Poisson distribution with mixing distribution $X \stackrel{\mathcal{L}}{=} \text{Po}(\lambda)$ given by an (ordinary) Poisson distribution with parameter λ , scaled by ρ :

$$\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} \sum_{m \geq 0} m^\ell e^{-\rho m} \left(e^{-\lambda} \frac{\lambda^m}{m!} \right) = \frac{\rho^\ell}{\ell!} e^{-\lambda + \lambda e^{-\rho}} \sum_{j=0}^{\ell} \left\{ \begin{matrix} \ell \\ j \end{matrix} \right\} (\lambda e^{-\rho})^j.$$

For a very comprehensive list of examples of mixed Poisson distributions we refer the reader to the article of Willmot [77]. Since by (1.3) the factorial moments $\mathbb{E}(Y^{\underline{s}})$ are related to the ordinary moments in terms of the Stirling numbers of the second kind, the moment sequence of Y is the (scaled) Stirling transform of the moment sequence of X . Next we collect similar basic properties of mixed Poisson distributions.

Proposition 1. Let $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ denote a mixed Poisson distributed random variable with mixing distribution X and scale parameter $\rho > 0$.

- (a) The factorial moments of Y are given by the scaled power moments of its mixing distribution, $\mathbb{E}(Y^{\underline{s}}) = \rho^s \mathbb{E}(X^s)$, $s \geq 1$.
- (b) The power moments of Y and X are related by the generalized Stirling transform with parameter ρ , and its inverse, respectively:

$$\mathbb{E}(Y^s) = \sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \rho^j \mathbb{E}(X^j), \quad \mathbb{E}(X^s) = \frac{1}{\rho^s} \sum_{j=0}^s (-1)^{s-j} \left[\begin{matrix} s \\ j \end{matrix} \right] \mathbb{E}(Y^j).$$

Similarly, the cumulants of Y and X are related by the generalized Stirling transform with parameter ρ , and its inverse, respectively.

- (c) The moment generating functions $\varphi(z) = \mathbb{E}(e^{zY})$ and $\psi(z) = \mathbb{E}(e^{zX})$ are related by the (generalized) Stirling transform of functions and its inverse, respectively:

$$\varphi(z) = \psi\left(\rho(e^z - 1)\right), \quad \psi(z) = \varphi\left(\log\left(1 + \frac{z}{\rho}\right)\right). \quad (2.3)$$

(d) Let $Y_1 \stackrel{\mathcal{L}}{=} MPo(\rho_1 X_1)$ and $Y_2 \stackrel{\mathcal{L}}{=} MPo(\rho_2 X_2)$ denote two independent mixed Poisson distributed random variables. Then, the sum $Y = Y_1 \oplus Y_2$ is again mixed Poisson distributed,

$$Y \stackrel{\mathcal{L}}{=} MPo(\rho_1 X_1 \oplus \rho_2 X_2).$$

Proof. (a) First we derive the factorial moments $\mathbb{E}(Y^{\underline{s}}) = \mathbb{E}(Y(Y-1)\dots(Y-s+1))$ of Y by a direct computation:

$$\begin{aligned} \mathbb{E}(Y^{\underline{s}}) &= \sum_{\ell \geq 0} \ell^{\underline{s}} \mathbb{P}\{Y = \ell\} = \sum_{\ell \geq s} \ell^{\underline{s}} \frac{\rho^\ell}{\ell!} \int_{\mathbb{R}^+} X^\ell e^{-\rho X} d\Lambda \\ &= \rho^s \sum_{\ell \geq s} \frac{\rho^{\ell-s}}{(\ell-s)!} \mathbb{E}(X^\ell e^{-\rho X}) = \rho^s \mathbb{E}\left(X^s e^{-\rho X} \sum_{\ell \geq 0} \frac{\rho^\ell X^\ell}{\ell!}\right) = \rho^s \mathbb{E}(X^s). \end{aligned}$$

(b) By converting $Y^{\underline{s}}$ into ordinary powers (1.3) the sequence of ordinary power moments $(\mathbb{E}(Y^s))_{s \in \mathbb{N}}$ of a mixed Poisson distributed random variable Y is given by the Stirling transform of the moments of the mixing distribution in the following way:

$$\mathbb{E}(Y^s) = \mathbb{E}\left(\sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} Y^j\right) = \sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \mathbb{E}(Y^j) = \sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \rho^j \mathbb{E}(X^j), \quad s \geq 1. \quad (2.4)$$

The result concerning the moment generating function in (c) can be shown similar to (1.4) by directly computing $\mathbb{E}(e^{zY})$, interchanging integration and summation:

$$\mathbb{E}(e^{zY}) = \sum_{\ell \geq 0} e^{z\ell} \mathbb{P}\{Y = \ell\} = \int_{\mathbb{R}^+} \sum_{\ell \geq 0} \frac{(e^z \rho X)^\ell}{\ell!} e^{-\rho X} d\Lambda = \int_{\mathbb{R}^+} e^{\rho(e^z - 1)X} d\Lambda.$$

By definition, the latter expression is exactly $\psi(\rho(e^z - 1))$, where $\psi(z) = \mathbb{E}(e^{zX})$ denotes the moment generating function of the mixing distribution X . If the cumulative distribution function of X is not known, we can compute the moment generating function $\varphi(z)$ of Y utilizing only the moment sequences:

$$\varphi(z) = \mathbb{E}(e^{zY}) = \sum_{s \geq 0} \mathbb{E}(Y^s) \frac{z^s}{s!} = \sum_{s \geq 0} \sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \rho^j \mu_j \frac{z^s}{s!} = \sum_{j \geq 0} \rho^j \mu_j \sum_{s \geq j} \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \frac{z^s}{s!}.$$

Using the bivariate generating function identity of the Stirling numbers of the second kind (see Wilf [76])

$$\sum_{s \geq 0} \sum_{j \geq 0} \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \frac{z^s}{s!} u^j = e^{u(e^z - 1)}, \quad (2.5)$$

we obtain further

$$\varphi(z) = \sum_{j \geq 0} \mu_j \frac{\rho^j (e^z - 1)^j}{j!}.$$

The latter expression is exactly the Stirling transform of $\psi(z) = \sum_{j \geq 0} \mu_j \frac{z^j}{j!}$ - in other words, of the moment generating function of X evaluated at $\rho(e^z - 1)$. The relation for the cumulants now follows readily from (c), since the cumulant generating functions $k_X(z)$ and $k_Y(z)$ of X and Y are given by $k_X(z) = \log(\psi(z))$ and $k_Y(z) = \log(\varphi(z))$. For a proof of part (d) we refer the reader to Johnson, Kotz and Kemp [41]. \square

In the applied probability literature, see [42, 77], given $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ it is usually assumed that the cumulative distribution function of the mixing distribution of X is known. However, in many cases in the analysis of random discrete structures the mixing distribution X is solely determined by the sequence of moments $\mathbb{E}(X^s) = \mu_s \in \mathbb{R}^+$, $s \geq 1$. Hence, it is beneficial to express the probability mass function of a mixed Poisson distributed random variable solely in terms of the moments of X , justifying (2.2). Note that for specific mixed Poisson distributions different simpler formulas may exist (compare with Corollary 2).

Proposition 2. *Let X denote a random variable with moment sequence given by $(\mu_s)_{s \in \mathbb{N}}$ such that $\psi(z) = \mathbb{E}(e^{zX})$ exists in a neighbourhood of zero, including the value $z = -\rho$. A random variable Y with factorial moments given by $\mathbb{E}(Y^s) = \rho^s \mu_s$ has a mixed Poisson distribution $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ with mixing distribution X and scale parameter $\rho > 0$, and the sequence of power moments of Y is the Stirling transform of the moment sequence $(\mu_s)_{s \in \mathbb{N}}$. The probability mass function of Y is given by*

$$\mathbb{P}\{Y = \ell\} = \sum_{s \geq \ell} (-1)^{s-\ell} \binom{s}{\ell} \mu_s \frac{\rho^s}{s!}, \quad \ell \geq 0.$$

Proof. By our assumption on the existence of $\psi(z)$ in a neighbourhood of zero, it follows that $\varphi(z)$ is also analytic around $z = 0$, and the random variable Y is uniquely determined by its (factorial) moments. Consequently, Y has a mixed Poisson distribution. Moreover, the probability mass function of Y is obtained by

$$\begin{aligned} \mathbb{P}\{Y = \ell\} &= \frac{\rho^\ell}{\ell!} \mathbb{E}(X^\ell e^{-\rho X}) = \frac{\rho^\ell}{\ell!} \left(D_z^\ell \psi(z) \right) \Big|_{z=-\rho} = \frac{\rho^\ell}{\ell!} \left(D_z^\ell \sum_{s \geq 0} z^s \frac{\mu_s}{s!} \right) \Big|_{z=-\rho} \\ &= \frac{\rho^\ell}{\ell!} \sum_{s \geq \ell} (-\rho)^{s-\ell} \frac{\mu_s}{(s-\ell)!} = \sum_{s \geq \ell} (-1)^{s-\ell} \binom{s}{\ell} \mu_s \frac{\rho^s}{s!}. \end{aligned} \tag{2.6}$$

Alternatively, the formula for the probability mass function can formally be

obtained directly from the definition

$$\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} \int_0^\infty X^\ell e^{-\rho X} d\Lambda = \frac{1}{\ell!} \int_0^\infty \sum_{s \geq 0} (-1)^s \frac{(\rho X)^{s+\ell}}{s!} d\Lambda.$$

Interchanging summation and integration also leads to the stated result. \square

2.3. The method of moments and basic limit laws

The method of moments is a classical way of deriving limit laws (see for example Hwang and Neininger [33] and the references therein). Given a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ one first derives asymptotic expansions of the power moments; assume that the moments satisfy the asymptotic expansion

$$\mathbb{E}(X_n^s) = \lambda_n^s \cdot \mu_s \cdot (1 + o(1)), \quad s \geq 1, \quad (2.7)$$

with λ_n denoting non-negative scale parameters. Then, one considers the scaled random variables $\frac{X_n}{\lambda_n}$, and tries to prove convergence in distribution of $\frac{X_n}{\lambda_n}$ by using the Fréchet-Shohat moment convergence theorem [55]: if the power moments of $\frac{X_n}{\lambda_n}$ converge to the moments $(\mu_s)_{s \in \mathbb{N}}$, and the moment sequence $(\mu_s)_{s \in \mathbb{N}}$ determines a unique non-degenerate distribution, then the random variable $\frac{X_n}{\lambda_n}$ converges in distribution to X . A well-known sufficient criterion for the uniqueness of the distribution of X is Carleman's condition: the distribution of X is uniquely determined if

$$\sum_{s \geq 1} (\mu_{2s})^{-\frac{1}{2s}} = \infty. \quad (2.8)$$

Note that (2.8) is satisfied, whenever $\mathbb{E}(e^{zX})$ exists in a neighbourhood of zero. We obtain the following new result concerning mixed Poisson distributions.

Lemma 1 (Uniqueness of mixed Poisson distributions). *The moments of a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$, with $\rho \geq 0$ and non-negative mixing distribution X , satisfy Carleman's criterion if and only if the moments of X do so. Moreover, the moment generating function $\psi(z) = \mathbb{E}(e^{zX})$ exists in a neighbourhood of zero, if and only if $\varphi(z) = \mathbb{E}(e^{zY})$ exists in a neighbourhood of zero.*

Proof. Note first that the second part follows directly from Proposition 1 part (c). Assume now that the moments of Y satisfy Carleman's condition. We observe that the moments $(\mu_s)_{s \in \mathbb{N}}$ of X are bounded by the scaled power moments of Y , $\mu_s \leq \frac{1}{\rho^s} \mathbb{E}(Y^s) = \sum_{j=0}^s \binom{s}{j} \rho^{s-j} \mu_j$. Consequently, the distribution of X is also uniquely determined by its moment sequence:

$$\sum_{s \geq 1} (\mu_{2s})^{-\frac{1}{2s}} \geq \rho \sum_{s \geq 1} (\mathbb{E}(Y^{2s}))^{-\frac{1}{2s}} = \infty.$$

Conversely, assume that the moments of X satisfy Carleman's condition:

$$\sum_{s \geq 1} (\mu_{2s})^{-\frac{1}{2s}} = \infty.$$

The s^{th} power moment of Y can be estimated using the s^{th} factorial moment of Y in the following way

$$\begin{aligned}\mathbb{E}(Y^s) &= \sum_{\ell \geq 0} \ell^s \mathbb{P}\{Y = \ell\} = \sum_{\ell=0}^{2s-1} \ell^s \mathbb{P}\{Y = \ell\} + \sum_{\ell \geq 2s} \ell^s \mathbb{P}\{Y = \ell\} \\ &\leq (2s)^s \cdot 1 + \sum_{\ell \geq 2s} 2^s \ell^s \mathbb{P}\{Y = \ell\} \leq (2s)^s + 2^s \mathbb{E}(Y^s).\end{aligned}$$

This implies that

$$\mathbb{E}(Y^s) \leq 2^s (s^s + \mathbb{E}(Y^s)) \leq 2^s (s^s + \rho^s \mu_s) \leq 4^s (1 + \rho)^s \cdot \max\{s^s, \mu_s\}.$$

Consequently,

$$(\mathbb{E}(Y^{2s}))^{-\frac{1}{2s}} \geq \frac{1}{4(1+\rho)} \cdot \min\left\{\frac{1}{2s}, (\mu_{2s})^{-\frac{1}{2s}}\right\},$$

such that

$$\sum_{s \geq 1} (\mathbb{E}(Y^{2s}))^{-\frac{1}{2s}} \geq \frac{1}{4(1+\rho)} \sum_{s \geq 1} \min\left\{\frac{1}{2s}, (\mu_{2s})^{-\frac{1}{2s}}\right\}.$$

By Hölder's inequality, the moments of $X \geq 0$ satisfy for $0 < r < s$ the inequality

$$\mathbb{E}(X^r) \leq (\mathbb{E}(X^s))^{\frac{r}{s}}.$$

Hence, for integers $0 < r < s$ we have

$$\mu_{2r}^{\frac{1}{2r}} \leq \mu_{2s}^{\frac{1}{2s}},$$

and sequence $(m_s)_{s \in \mathbb{N}}$, defined by $m_s := \mu_{2s}^{-\frac{1}{2s}}$, is monotonically decreasing. It remains to show that

$$\sum_{s \geq 1} \min\left\{\frac{1}{s}, m_s\right\} = \infty, \quad (2.9)$$

which immediately implies the required result; note that we omitted the additional factor $\frac{1}{2}$ for the sake of simplicity. If m_s is bounded away from zero this is immediately true. Hence, in the following we assume that $(m_s)_{s \in \mathbb{N}}$ is a null sequence. Let $\mathbb{N} = I_1 \cup I_2$, with $I_1 \cap I_2 = \emptyset$, such that for all $s \in I_1$ we have $\frac{1}{s} \leq m_s$, and for $s \in I_2$ we have $\frac{1}{s} > m_s$. We obtain

$$\sum_{s \geq 1} \min\left\{\frac{1}{s}, m_s\right\} = \sum_{s \in I_1} \frac{1}{s} + \sum_{s \in I_2} m_s.$$

By our initial assumption $\sum_{s \geq 1} m_s = \infty$, equation (2.9) is directly satisfied if either I_1 or I_2 is finite. Hence, we assume that both sets are infinite. Assume

further that $\sum_{s \in I_1} \frac{1}{s}$ is finite. We can write the set I_1 as the disjoint union of infinitely many finite length intervals

$$I_1 = \bigcup_{\ell \geq 1} [a_\ell, b_\ell],$$

with $[a_\ell, b_\ell] := \{a_\ell, a_\ell + 1, \dots, b_\ell\}$ and $a_\ell, b_\ell \in \mathbb{N}$ for all $\ell \in \mathbb{N}$. If all but finitely many intervals are of length one, such that $a_\ell = b_\ell$, the values s with $\min\{\frac{1}{s}, m_s\} = \frac{1}{s}$ are essentially isolated. In this case we note that $\ell \in I_1$ and $\ell - 1 \in I_2$ and use for $\ell \geq 2$ the inequality

$$a_\ell \leq a_{\ell-1} \leq \frac{1}{\ell-1} \leq \frac{2}{\ell}.$$

This implies that also $\sum_{s \in I_1} m_s$ is finite too, such that $\sum_{s \in I_2} m_s$ is infinite. Finally, we assume that infinitely many intervals are of length greater or equal two. By our earlier assumption $\sum_{s \in I_1} \frac{1}{s}$ is finite and satisfies

$$\sum_{s \in I_1} \frac{1}{s} = \sum_{\ell \in \mathbb{N}} \sum_{s \in [a_\ell, b_\ell]} \frac{1}{s} > \sum_{\ell \in \mathbb{N}} \int_{a_\ell}^{b_\ell} \frac{1}{x} dx = \sum_{\ell \in \mathbb{N}} \ln \left(\frac{b_\ell}{a_\ell} \right) > 0.$$

Furthermore, $\ln \left(\frac{b_\ell}{a_\ell} \right) < \epsilon$ for all sufficiently large ℓ , such that $b_\ell < e^\epsilon a_\ell$. This implies that for $k \in [a_\ell, b_\ell]$ and sufficiently large ℓ

$$\frac{m_k}{\frac{1}{k}} \leq \frac{m_{a_\ell}}{\frac{1}{b_\ell}} < \frac{m_{a_\ell}}{\frac{1}{e^\epsilon a_\ell}} \leq \frac{m_{a_\ell-1}}{\frac{1}{e^\epsilon a_\ell}} \leq e^\epsilon \frac{1}{\frac{1}{a_\ell}} \leq 2e^\epsilon.$$

Hence, $m_k \leq \frac{2e^\epsilon}{k}$. Combining this with our previous argument for the essentially isolated values we deduce that $\sum_{s \in I_1} m_s$ is finite too, such that $\sum_{s \in I_2} m_s = \infty$. \square

Concerning random discrete structures one usually encounters non-negative discrete random variables X_n . As mentioned before in (2.1) the factorial moments are readily obtained from the probability generating function $p(z) = \mathbb{E}(z^{X_n})$ by repeated differentiation:

$$\mathbb{E}(X_n^{\underline{s}}) = E_z D_z^s p(z), \quad s \geq 1.$$

In contrast, the ordinary moments require the usage of the so-called theta differential operator $\Theta_z = zD_z$: $\mathbb{E}(X_n^s) = E_z \Theta_z^s p(z)$. Mixed Poisson distributions and a related phase transition naturally occur if the factorial moments satisfy asymptotic expansions similar to (2.7) instead of the power moments. We present a new characterization based on factorial moments.

Lemma 2 (Factorial moments and limit laws of mixed Poisson type). *Let $(X_n)_{n \in \mathbb{N}}$ denote a sequence of random variables, whose factorial moments are*

asymptotically of mixed Poisson type satisfying for n tending to infinity the asymptotic expansion

$$\mathbb{E}(X_n^s) = \lambda_n^s \cdot \mu_s \cdot (1 + o(1)), \quad s \geq 1,$$

with $\mu_s \geq 0$, and $\lambda_n > 0$. Furthermore assume that the moment sequence $(\mu_s)_{s \in \mathbb{N}}$ determines a unique distribution X satisfying Carleman's condition. Then, the following limit distribution results hold:

- (i) if $\lambda_n \rightarrow \infty$, for $n \rightarrow \infty$, the random variable $\frac{X_n}{\lambda_n}$ converges in distribution, with convergence of all moments, to X .
- (ii) if $\lambda_n \rightarrow \rho \in (0, \infty)$, for $n \rightarrow \infty$, the random variable X_n converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$.

Moreover, the random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ converges for $\rho \rightarrow \infty$, after scaling, to its mixing distribution $X: \frac{Y}{\rho} \stackrel{\mathcal{L}}{\rightarrow} X$, with convergence of all moments.

Remark 1. It may be possible to unify cases (i) and (ii) to arbitrary sequences λ_n by a suitable result for the distance between random variables X_n and $Y_n = \text{MPo}(\lambda_n X)$.

Remark 2. The results above complement the standard case when the distribution of X degenerates, $X = 1$. The random variables X_n are then asymptotically Poisson distributed with parameter λ_n . Thus, the distribution of $\frac{X_n}{\lambda_n}$ degenerates for $\lambda_n \rightarrow \infty$, since we expect a central limit theorem for $(X_n - \lambda_n)/\sqrt{\lambda_n}$. It might also be necessary for non-degenerate X to consider centered random variables similar to $X_n^* = X_n - \lambda_n$, and its (factorial) moments, instead of X_n .

Remark 3. The result above can be strengthened to also include the degenerate case $\lambda_n \rightarrow 0$, such that $X_n \stackrel{\mathcal{L}}{\rightarrow} 0$. By Markov's inequality it suffices to prove that $\mathbb{E}(X_n) \rightarrow 0$. In order to obtain additional moment convergence one has to show $\mathbb{E}(X_n^s) \rightarrow 0$ for every $s \in \mathbb{N}$.

Remark 4 (Moment generating functions and limit laws of mixed Poisson type). Let $\psi(z) = \mathbb{E}(e^{zX})$ denote the moment generating function of X . If the moment generating function $\varphi(z) = \mathbb{E}(e^{zX_n})$ satisfies for $n \rightarrow \infty$ the asymptotic expansion

$$\varphi(z) = \psi(\lambda_n(e^z - 1)) \cdot (1 + o(1)),$$

then the conclusion of the lemma above - convergence in distribution - still holds, but a priori without moment convergence. On the other hand, if the moments of $(\mu_s)_{s \in \mathbb{N}}$ do not determine a unique distribution, one still obtains by the Lemma above convergence of integer moments, but one cannot deduce convergence in distribution.

Remark 5. In the analysis of random discrete structures the random variables X_n often depend on an additional parameter describing or measuring a certain local aspect of the combinatorial structure, such that $X_n = X_{n,j}$. Moreover, the

expansion of the factorial moments often depend on this parameter in a crucial way. A quite common situation (see [47, 49, 53, 65] and also [21, 37]) is the following dichotomy for the asymptotic expansion of the factorial moments:

$$\mathbb{E}(X_n^s) = \begin{cases} \lambda_{n,0}^s \cdot \mu_{s,j} \cdot (1 + o(1)), & s \geq 1, \quad j \text{ fixed,} \\ \lambda_{n,1}^s \cdot \mu_s \cdot (1 + o(1)), & s \geq 1, \quad j \rightarrow \infty, \end{cases}$$

where $\lambda_{n,0}$ is independent of j , but $\lambda_{n,1} = \lambda_{n,1}(j)$ also depends on the growth of this additional parameter j compared to n . Consequently, one encounters one additional family of limit laws when j is fixed, determined by the moment sequence $(\mu_{s,j})_{s \in \mathbb{N}}$. Note that in all presented examples the following additional property holds for $s \geq 1$:

$$\Lambda_j^s \mu_{s,j} \rightarrow \mu_s, \quad j \rightarrow \infty,$$

where Λ_j denotes an additional scale parameter; compare with the Remarks 6, 7, and 13.

Proof of Lemma 2. By (1.3) the power moments of X_n satisfy the following asymptotic expansion

$$\mathbb{E}(X_n^s) = \sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \mathbb{E}(X_n^j) = \sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \lambda_n^j \mu_j (1 + o(1)) = \left(\sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \lambda_n^j \mu_j \right) (1 + o(1)).$$

If $\lambda_n \rightarrow \infty$ for $n \rightarrow \infty$, we obtain further the expansion

$$\begin{aligned} \mathbb{E}(X_n^s) &= \left(\sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \lambda_n^j \mu_j \right) (1 + o(1)) = \left(\left\{ \begin{matrix} s \\ s \end{matrix} \right\} \lambda_n^s \mu_s + \mathcal{O}(\lambda_n^{s-1}) \right) (1 + o(1)) \\ &= \lambda_n^s \mu_s + \mathcal{O}(\lambda_n^{s-1}) + o(\lambda_n^s). \end{aligned}$$

Consequently, the moments of $\frac{X_n}{\lambda_n}$ converge to the moments μ_s of the mixing distribution. Since the moments of X satisfy Carleman's condition, this proves by the Fréchet-Shohat moment convergence theorem convergence in distribution.

Furthermore, for $\lambda_n \rightarrow \rho$ for $n \rightarrow \infty$, we directly obtain

$$\mathbb{E}(X_n^s) = \left(\sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \lambda_n^j \mu_j \right) (1 + o(1)) = \sum_{j=0}^s \left\{ \begin{matrix} s \\ j \end{matrix} \right\} \rho^j \mu_j + o(1).$$

Consequently, the moments of X_n converge to the moments of a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$, which is uniquely determined by its moment sequence, according to Lemma 1, and our assumption on the moments of X . Finally, an identical argument proves that a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ converges to its mixing distribution for $\rho \rightarrow \infty$. \square

3. Examples and applications

We present several appearances of mixed Poisson distributions in the analysis of random discrete structures, in particular various families of random trees, k -Stirling permutations, and urn models. We discuss several families of random trees where a mixed Poisson law arises as the limit law of a discrete random variable $X_{n,j}$. The parameter $n \in \mathbb{N}$ usually measures the size of the investigated trees, and j denotes an additional parameter measuring or marking a certain aspect of the combinatorial structure, i.e., a node with a certain label j of interest, often satisfying a natural constraint of the type $1 \leq j \leq n$, see [47, 49, 51, 64]. In the limit $n \rightarrow \infty$, with $j = j(n)$, phase transitions were observed according to the relative growth of j with respect to n , e.g., $j = 1, 2, \dots$ being a constant independent of n , $j \rightarrow \infty$ but with $j = o(n)$, or $j \sim \rho \cdot n$, for fixed ρ . As mentioned in the introduction, in this section we will unify and simplify earlier arguments by starting from explicit formulas for the factorial moments occurring in various works. These explicit formulas directly lead to mixed Poisson laws, using Lemmas 1 and 2, and Stirling's formula for the Gamma function

$$\Gamma(x) = x^{x-\frac{1}{2}} e^{-x} \sqrt{2\pi} (1 + \mathcal{O}(x^{-1})), \quad \text{for } x \rightarrow \infty. \quad (3.1)$$

Besides that, whenever possible we give an interpretation of the random variables occurring in terms of urn models.

3.1. Block sizes in k -Stirling permutations

Stirling permutations were defined by Gessel and Stanley [29]. A Stirling permutation is a permutation of the multiset $\{1, 1, 2, 2, \dots, n, n\}$ such that, for each i , $1 \leq i \leq n$, the elements occurring between the two occurrences of i are larger than i . E.g., 1122, 1221 and 2211 are Stirling permutations, whereas the permutations 1212 and 2112 of $\{1, 1, 2, 2\}$ aren't. The name of these combinatorial objects is due to relations with the Stirling numbers, see [29] for details and [44] for bijections with certain tree families. A straightforward generalization of Stirling permutations is to consider permutations of a more general multiset $\{1^k, 2^k, \dots, n^k\}$, with $k \in \mathbb{N}$ (we use in this context $j^\ell := \underbrace{j, \dots, j}_\ell$, for $\ell \geq 1$),

such that for each i , $1 \leq i \leq n$, the elements occurring between two occurrences of i are at least i . Such restricted permutations on the multiset $\{1^k, 2^k, \dots, n^k\}$ are called k -Stirling permutations of order n ; they have already been considered by Brenti [13, 14], Park [66, 67, 68], and Janson et al. [38, 40]. These k -Stirling permutations can be generated in a sequential manner: we start with $1^k = 1 \dots 1$ and insert the string $(n+1)^k$ at any position (anywhere, including first or last) in a given k -Stirling permutation of $\{1^k, 2^k, \dots, n^k\}$, $n \geq 1$. In the case $k = 3$, we have for example one permutation of order 1: 111; four permutations of order 2: 111222, 112221, 122211, 222111; etc.

A *block* in a k -Stirling permutation $\sigma = \sigma_1 \dots \sigma_s$ is a substring $\sigma_a \dots \sigma_b$, with $\sigma_a = \sigma_b$, that is maximal, i.e., which is not contained in any larger such

substring. There is obviously at most one block for every $j \in \{1, 2, \dots, n\}$, extending from the first occurrence of j to the last one; we say that j forms a block if this substring is indeed a block, i.e., when it is not contained in a string $j' \cdots j'$, for some $j' < j$. It can be shown easily by induction that any k -Stirling permutation has a unique decomposition as a sequence of its blocks. For example, the 3-Stirling permutation $\sigma = 1122333214455547777666$, has block decomposition $[112233321][445554][777][666]$.

Of course, the size of a block in a k -Stirling permutation is always a multiple of k . The number of blocks of size $k \cdot \ell$ in a random k -Stirling permutation of order n was studied in [51]. There, a simple exact expression for the factorial moments was derived:

$$\mathbb{E}(X_{n,\ell}^s) = \frac{s!}{(k\ell)^s} \binom{\ell - 1 - \frac{1}{k}}{\ell - 1}^s \cdot \frac{\binom{n - \ell s + \frac{s+1}{k} - 1}{n - \ell s}}{\binom{n - 1 + \frac{1}{k}}{n}}.$$

Depending on the growth of $\ell = \ell(n)$ as $n \rightarrow \infty$, two random variables X and Y arose as limiting distributions of $X_{n,\ell}$. The random variable X with moment sequence

$$\mathbb{E}(X^s) = (s+1)! \frac{\Gamma(1 + \frac{1}{k})}{\Gamma(1 + \frac{s+1}{k})} \quad (3.2)$$

could be characterized using observations by Janson et al. [40], and Janson [37]. It has a density function $f(x)$ that can be written as

$$f(x) = \frac{\Gamma(\frac{1}{k})}{\pi} \sum_{j=1}^{\infty} (-1)^{j-1} \frac{\Gamma(\frac{j}{k} + 1) \sin \frac{j\pi}{k}}{j!} x^j, \quad \text{for } x > 0. \quad (3.3)$$

However, the characterization of the random variable Y was incomplete, only the (factorial) moments were known:

$$\mathbb{E}(Y^s) = (s+1)! \frac{\Gamma(1 + \frac{1}{k})}{\Gamma(1 + \frac{s+1}{k})} \rho^s, \quad s \geq 1. \quad (3.4)$$

Using Lemma 2, we can fill this gap, extending the results of [51].

Corollary 1. *The factorial moments of the random variable $X_{n,\ell}$, counting the number of blocks of size $k \cdot \ell$ in a random k -Stirling permutation of order n , are for $n \rightarrow \infty$ asymptotically of mixed Poisson type, with mixing distribution X , determined by its moments and density given by (3.2) and (3.3) and scale parameter $\lambda_{n,\ell} = \frac{1}{k\ell} \binom{\ell - 1 - \frac{1}{k}}{\ell - 1} n^{\frac{1}{k}}$:*

$$\mathbb{E}(X_{n,\ell}^s) = \lambda_{n,\ell}^s (s+1)! \frac{\Gamma(1 + \frac{1}{k})}{\Gamma(1 + \frac{s+1}{k})} (1 + o(1)),$$

(i) for $\ell = \ell(n)$, such that $\lambda_{n,\ell} \rightarrow \infty$, the random variable $\frac{X_{n,\ell}}{\lambda_{n,\ell}}$ converges in distribution, with convergence of all moments, to X .

(ii) for $\ell = \ell(n)$, such that $\lambda_{n,\ell} \rightarrow \rho \in (0, \infty)$, the random variable $X_{n,\ell}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$. Its probability mass function is given by

$$\mathbb{P}\{Y = i\} = \sum_{s \geq i} \binom{s}{i} (-1)^{s-i} \rho^s \frac{\Gamma(1 + \frac{1}{k})}{\Gamma(1 + \frac{s+1}{k})}, \quad i \geq 0.$$

Moreover, for $\rho \rightarrow \infty$, the random variable Y/ρ converges in distribution to X , with convergence of all moments.

The result above can also be interpreted in terms of a suitable urn model. First we recall the definition of Pólya-Eggenberger urn models. We start with an urn containing n white balls and m black balls. The evolution of the urn occurs in discrete time steps. At every step a ball is drawn at random from the urn. The color of the ball is inspected and then the ball is returned to the urn. According to the observed color of the ball there are added/removed balls due to the following rules. If a white ball has been drawn, we put into the urn α white balls and β black balls, but if a black ball has been drawn, we put into the urn γ white balls and δ black balls. The values $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ are fixed integer values and the urn model is specified by the 2×2 ball replacement matrix $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. This definition readily extends to higher dimensions, leading to $r \times r$ ball replacement matrices, if balls of r different colours are involved. Note that we may consider $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, defining the urn process as a Markov process; see Remark 1.11 of Janson [37]. One usually assumes that the urns are tenable: the process of drawing and adding/removing balls can be continued ad infinitum, never having to remove balls which are not present in the urn. Starting with $W_0 = w_0$ white balls and $B_0 = b_0$ black balls, one is then interested in the composition (W_n, B_n) of the urn after n draws. For a few recent results we refer the reader to [10, 21, 22, 34, 37, 70].

In order to describe the growth of the r.v. $X_{n,i}$, $1 \leq i \leq \ell$, by means of a Pólya-Eggenberger urn model, we consider the simple growth process of random k -Stirling permutations: in order to generate a random k -Stirling permutation of order $n+1$, we select uniformly at random a k -Stirling permutation of order n and insert the substring $(n+1)^k$ uniformly at random at one of the $kn+1$ insertion positions. In the urn model description, each ball in the urn will correspond to an insertion position in the k -Stirling permutation. We will require $\ell+2$ different colours of balls. Balls of colours i , with $1 \leq i \leq \ell$, correspond to insertion positions within blocks of size ki , balls of colour $\ell+1$ correspond to insertion positions within blocks of size $\geq k(\ell+1)$, and balls of colour 0 correspond to insertion positions between two consecutive blocks, or before the first or after the last block. When inserting a new substring the following situations can occur, which then describe the evolution of the urn:

- (i) When inserting the string $(n+1)^k$ into a block of size ki , $1 \leq i \leq \ell$, then this block changes to a block of size $k(i+1)$. In the urn model this means

that if a ball of colour i is drawn, $ki - 1$ balls of colour i have to be removed and $k(i + 1) - 1$ balls of colour $i + 1$ have to be added.

- (ii) When inserting the string $(n + 1)^k$ into a block of size $\geq k(\ell + 1)$, then it remains a block of size $\geq k(\ell + 1)$, but its size increases by k . In the urn model this means that if a ball of colour $\ell + 1$ is drawn, k balls of colour $\ell + 1$ have to be added.
- (iii) When inserting the string $(n + 1)^k$ between two consecutive blocks, or before the first or after the last block, then a new block of size k appears; furthermore an additional insertion position between two consecutive blocks occurs. In the urn model this means that if a ball of colour 0 is drawn, $k - 1$ balls of colour 1 and one ball of colour 0 have to be added.

The initial k -Stirling permutation 1^k contains one block of size $k \cdot 1$ with $k - 1$ insertion positions; furthermore there are one insertion position before the first and one insertion position after the last block, which describes the initial configuration of the urn. Thus the following urn model description follows immediately.

Urn I. Consider a balanced urn (i.e., each row of the ball replacement matrix has the same row sum) with balls of $\ell + 2$ colours and let the random vector $(U_{n,0}, \dots, U_{n,\ell+1})$ count the number of balls of each colour at time n with the $(\ell + 2) \times (\ell + 2)$ ball replacement matrix M given by

$$M = \begin{pmatrix} 1 & k-1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & -(k-1) & 2k-1 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & -(2k-1) & 3k-1 & \ddots & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & -((\ell-1)k-1) & \ell k-1 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 & -(\ell k-1) & (\ell+1)k-1 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & k \end{pmatrix}.$$

The initial configuration of the urn (it is here convenient to start at time 1) is given by $(U_{1,0}, \dots, U_{1,\ell+1}) = (2, k - 1, 0, \dots, 0)$. It holds that the random variables $U_{n,i}$, with $1 \leq i \leq \ell$, described by the urn model are related to the random variables $X_{n,i}$, $1 \leq i \leq \ell$, which count the number of blocks of size ki in a random k -Stirling permutation of order n , as follows:

$$U_{n,i} = (ki - 1)X_{n,i}, \quad 1 \leq i \leq \ell.$$

By Theorem 1 and the results of [51] this implies that the random variables $U_{n,i}$ occurring in the urn model undergo a phase transition according to the growth of ℓ with respect to n , from continuous to discrete, where the moments of the appearing random variables X and Y are related by the Stirling transform.

3.2. Diminishing Pólya-Eggenberger urn models

A classical example of a non-tenable urn model is the sampling without replacement urn with ball replacement matrix given by $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. The process of drawing

and replacing balls ends after $n + m$ steps, starting with n white and m black balls. Here, one is interested in the number of white balls remaining in the urn, after all black balls have been drawn. Several urn models of a similar non-tenable nature have recently received some attention under the name diminishing urn models, see [53] and the references therein.

Urn II. Consider a possibly unbalanced generalized sampling without replacement urn model with ball replacement matrix

$$M = \begin{pmatrix} -\alpha & 0 \\ 0 & -\delta \end{pmatrix}, \quad a, \delta \in \mathbb{N}.$$

The initial configuration of the urn consists of $\alpha \cdot n$ white balls and $\delta \cdot m$ black balls. The random variable $X_{\delta m, \alpha n}$ counts the number of white balls remaining in the urn, after all black balls have been drawn.

It was shown in [53] that the factorial moments of the random variable $\hat{X}_{\delta m, \alpha n} = X_{\delta m, \alpha n} / \alpha$ are given by

$$\mathbb{E}\left(\hat{X}_{\delta m, \alpha n}^s\right) = \frac{n^s}{\binom{m + \frac{\alpha s}{\delta}}{m}}, \quad s \geq 1.$$

Moreover, a random variable Y arises in the limit, whose factorial moments are given by

$$\mathbb{E}(Y^s) = \rho^s \Gamma\left(1 + \frac{\alpha s}{\delta}\right), \quad s \geq 1. \quad (3.5)$$

Using a special case of Theorem 2 it was shown that Y has a discrete distribution. However, the result of [53] contains a small gap: the moments $(\Gamma(1 + \frac{\alpha s}{\delta}))_{s \in \mathbb{N}}$ only determine a unique distribution for $\alpha/\delta \leq 2$, see [32]. Hence, only in this case the (factorial) moments of Y determine a unique distribution. Since a Weibull distributed random variable $X \stackrel{\mathcal{L}}{=} W_{\delta/\alpha, 1}$, with shape parameter $\frac{\delta}{\alpha}$, scale parameter 1, and density $f(t) = \frac{\delta}{\alpha} t^{\frac{\delta}{\alpha}-1} e^{-t^{\frac{\delta}{\alpha}}}$, $t \geq 0$, has moments $\mathbb{E}(X^s) = \Gamma(1 + \frac{\alpha s}{\delta})$, we obtain the following characterization of Y , extending the result of [53].

Corollary 2. *The random variable $\hat{X}_{\delta m, \alpha n}$, counting the number of white balls remaining in the urn, after all black balls have been drawn in a generalized sampling without replacement urn, starting with $\alpha \cdot n$ white balls and $\delta \cdot m$ black balls, has for $\min\{n, m\} \rightarrow \infty$ factorial moments of mixed Poisson type with a Weibull mixing distribution $X \stackrel{\mathcal{L}}{=} W_{\delta/\alpha, 1}$, and scale parameter $\lambda_{n, m} = \frac{n}{m \frac{\alpha}{\delta}}$:*

$$\mathbb{E}\left(\hat{X}_{\delta m, \alpha n}^s\right) = \lambda_{n, m}^s \Gamma\left(1 + \frac{\alpha s}{\delta}\right)(1 + o(1)).$$

Assume that $\alpha/\delta \leq 2$:

- (i) for $\lambda_{n, m} \rightarrow \infty$, the random variable $\frac{\hat{X}_{\delta m, \alpha n}}{\lambda_{n, m}}$ converges in distribution, with convergence of all moments, to X .

(ii) for $\lambda_{n,m} \rightarrow \rho \in (0, \infty)$, the random variable $\hat{X}_{\delta m, \alpha n}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$. Its probability mass function is given as follows:

$$\mathbb{P}\{Y = \ell\} = \begin{cases} \sum_{j \geq \ell} (-1)^{j-\ell} \binom{j}{\ell} \rho^j \frac{\Gamma(1 + \frac{\alpha j}{\delta})}{j!}, & \text{for } \frac{\alpha}{\delta} < 1, \\ \frac{1}{1 + \rho} \left(\frac{\rho}{1 + \rho} \right)^\ell, & \text{for } \frac{\alpha}{\delta} = 1, \\ \frac{\delta}{\alpha} \sum_{j \geq 0} (-1)^j \binom{j + \ell}{\ell} \rho^{-\frac{\delta}{\alpha}(j+1)} \frac{\Gamma(\frac{\delta(j+1)}{\alpha} + \ell)}{(j + \ell)!}, & \text{for } \frac{\alpha}{\delta} > 1. \end{cases}$$

Remark 6. As shown in [53], for fixed m the random variable $\hat{X}_{\delta m, \alpha n}/n$ converges to the power $B^{\frac{1}{\alpha}}$ of a beta-distributed random variable B , with moments $\mathbb{E}(B^s) = 1/\binom{m + \frac{\alpha s}{\delta}}{m}$. The Weibull mixing distribution $X \stackrel{\mathcal{L}}{=} W_{\delta/\alpha, 1}$ can be recovered by considering the limit $m \rightarrow \infty$ of $B = B_m$:

$$m^{\frac{\alpha}{\delta}} B_m \rightarrow X, \quad \text{for } m \rightarrow \infty,$$

with convergence of all moments. Note that the results above can be extended to all $\alpha, \delta \in \mathbb{N}$; however, for $\alpha/\delta \geq 2$ the method of moments cannot be used anymore. Instead, has to directly analyse the probability generating function $h_{n,m}(v)$, which can be derived using stochastic processes [52].

Proof of Corollary 2. According to the definition of a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$, it has factorial moments given by (3.5). In order to derive the integral-free series representation we proceed as follows. In the first case $\alpha/\delta < 1$ we can directly use Theorem 2, since the moment generating function of the mixing Weibull distribution X exists at $-\rho$. In the remaining cases $\alpha/\delta \geq 1$ we use the definition and the density function of the Weibull distribution to get first

$$\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} \int_0^\infty \frac{\delta}{\alpha} t^{\frac{\delta}{\alpha} + \ell - 1} \exp\left(-t^{\frac{\delta}{\alpha}} - \rho t\right) dt.$$

The case $\alpha/\delta = 1$ readily leads to the stated geometric distribution after using the obvious simplification

$$\frac{\delta}{\alpha} t^{\frac{\delta}{\alpha} + \ell - 1} \exp\left(-t^{\frac{\delta}{\alpha}} - \rho t\right) = t^\ell e^{-t(\rho+1)}.$$

Finally, for $\alpha/\delta > 1$ we expand $e^{-t^{\frac{\delta}{\alpha}}} = \sum_{j \geq 0} (-1)^j \frac{t^{\frac{j\delta}{\alpha}}}{j!}$ and obtain

$$\mathbb{P}\{Y = \ell\} = \frac{\delta}{\alpha \cdot \ell!} \sum_{j \geq 0} \frac{(-1)^j}{j!} \int_0^\infty t^{\frac{(j+1)\delta}{\alpha} + \ell - 1} e^{-\rho t} dt.$$

The Gamma-function type integrals are readily evaluated and the stated result follows. \square

3.3. Descendants in increasing trees.

Increasing trees are labelled trees, where the nodes of a tree of size n are labelled by distinct integers of the set $\{1, \dots, n\}$ in such a way that each sequence of labels along any branch starting at the root is increasing. They have been introduced by Bergeron et al. [11], and can be described combinatorially as follows. Given a so-called degree-weight sequence $(\varphi_k)_{k \geq 0}$, the corresponding degree-weight generating function $\varphi(t)$ is defined by $\varphi(t) := \sum_{k \geq 0} \varphi_k t^k$. The simple family of increasing trees \mathcal{T} associated with a degree-weight generating function $\varphi(t)$, can be described by the formal recursive equation

$$\mathcal{T} = \textcircled{1} \times \left(\varphi_0 \cdot \{\epsilon\} \dot{\cup} \varphi_1 \cdot \mathcal{T} \dot{\cup} \varphi_2 \cdot \mathcal{T} * \mathcal{T} \dot{\cup} \varphi_3 \cdot \mathcal{T} * \mathcal{T} * \mathcal{T} \dot{\cup} \dots \right) = \textcircled{1} \times \varphi(\mathcal{T}), \quad (3.6)$$

where $\textcircled{1}$ denotes the node labelled by 1, \times the Cartesian product, $\dot{\cup}$ the disjoint union, $*$ the partition product for labelled objects, and $\varphi(\mathcal{T})$ the substituted structure (see, e.g., the books [24, 28]). Note that the elements of \mathcal{T} are increasing plane trees, and that such a tree of size n , whose nodes have outdegrees d_1, \dots, d_n , has the weight $\prod_{i=1}^n \varphi_{d_i}$. When speaking about a random tree of size n from the family \mathcal{T} , we always use the random tree model for weighted trees, i.e., we assume that each tree of size n from \mathcal{T} can be chosen with a probability proportional to its weight.

Let T_n be the total weight of all trees from \mathcal{T} of size n . It follows from (3.6) that the exponential generating function $T(z) := \sum_{n \geq 1} T_n \frac{z^n}{n!}$ of the total weights satisfies the autonomous first order differential equation

$$T'(z) = \varphi(T(z)), \quad T(0) = 0. \quad (3.7)$$

From now on we consider tree families \mathcal{T} having degree-weights of one of the following three forms, as studied by [63], where we use the abbreviations RECT for recursive trees, GPORT for generalized plane recursive trees (also called generalized plane-oriented recursive trees), and d -INCT for d -ary increasing trees.

$$\varphi(t) = \begin{cases} e^{c_1 t}, & \text{for } c_1 > 0, \quad \text{for RECT,} \\ \frac{\varphi_0}{\left(1 + \frac{c_2 t}{\varphi_0}\right)^{\frac{c_1}{c_2} - 1}}, & \text{for } \varphi_0 > 0, 0 < -c_2 < c_1, \quad \text{for GPORT,} \\ \varphi_0 \left(1 + \frac{c_2 t}{\varphi_0}\right)^d, & \text{for } \varphi_0, c_2 > 0, d := \frac{c_1}{c_2} + 1 \in \mathbb{N} \setminus \{1\}, \text{ for } d\text{-INCT.} \end{cases} \quad (3.8)$$

Consequently, by solving (3.7), we obtain the exponential generating function $T(z)$,

$$T(z) = \begin{cases} \log\left(\frac{1}{1-c_1 z}\right), & \text{for RECT,} \\ \frac{\varphi_0}{c_2} \left(\frac{1}{(1-c_1 z)^{\frac{c_2}{c_1}}} - 1\right), & \text{for GPORT,} \\ \frac{\varphi_0}{c_2} \left(\frac{1}{(1-(d-1)c_2 z)^{\frac{1}{d-1}}} - 1\right), & \text{for } d\text{-INCT,} \end{cases} \quad (3.9)$$

and the total weights T_n ,

$$T_n = \varphi_0 c_1^{n-1} (n-1)! \binom{n-1 + \frac{c_2}{c_1}}{n-1}. \quad (3.10)$$

Note that changing φ_k to $ab^k\varphi_k$ for some positive constants a and b will affect the weights of all trees of a given size n by the same factor $a^n b^{n-1}$, which does not affect the distribution of a random tree from the family. Hence, when considering random trees from these three classes, φ_0 is irrelevant and c_1 and c_2 are relevant only through the ratio c_1/c_2 . (We may thus, if we like, normalize $\varphi_0 = 1$ and either c_1 or $|c_2|$, but not both.) It is convenient to set $c_1 = 1$ for (random) recursive trees, to use the parameter $\alpha := -1 - \frac{c_1}{c_2} > 0$ for (random) generalized plane recursive trees, and $d := \frac{c_1}{c_2} + 1 \in 2, 3, \dots$ for (random) d -ary increasing trees, i.e., it suffices to consider the degree-weight generating functions

$$\varphi(t) = \begin{cases} e^t, & \text{for RECT,} \\ \frac{1}{(1-t)^\alpha}, & \text{with } \alpha > 0, \text{ for GPORT,} \\ (1+t)^d, & \text{with } d = 2, 3, 4, \dots, \text{ for } d\text{-INCT.} \end{cases} \quad (3.11)$$

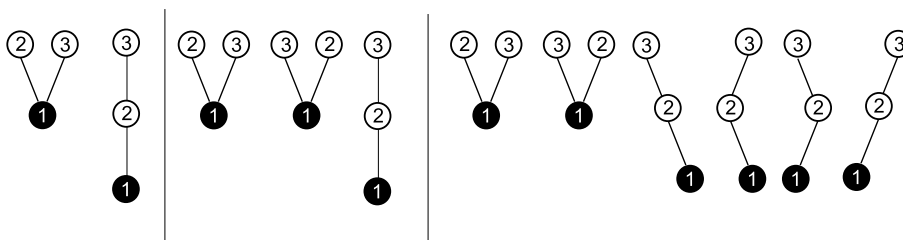


FIG 1. Two recursive trees of size three - $\varphi(t) = e^t$ (no left-to-right order); three plane recursive trees of size three - $\varphi(t) = \frac{1}{1-t}$; six binary increasing trees - $\varphi(t) = (1+t)^2$.

As shown by Panholzer and Prodinger [63], random trees in the three classes of families given in (3.11) can be generated by an evolution process in the following way. The process, evolving in discrete time, starts with the root labelled by 1. At step $i+1$ the node with label $i+1$ is attached to any previous node v (with outdegree $d(v)$) of the already grown tree of size i with probabilities $p(v)$ given by

$$p(v) = \begin{cases} \frac{1}{i}, & \text{for RECT,} \\ \frac{d(v)+\alpha}{(\alpha+1)^{i-1}}, & \text{for GPORT,} \\ \frac{d-d(v)}{(d-1)^{i+1}}, & \text{for } d\text{-INCT.} \end{cases} \quad (3.12)$$

Moreover, it has been shown in [63] that only random trees from simple families of trees \mathcal{T} given in (3.8) can be generated by such a tree evolution process, i.e., only for these tree families exist suitable attachment probabilities $p(v)$.

Let $D_{n,j}$ denote the random variable, counting the number of descendants, i.e., the size of the subtree rooted at node j , of a specific node j , with $1 \leq j \leq n$, in a tree of size n . In [47] this random variable has been studied for the three tree families mentioned beforehand using a generating functions approach. In the following we collect and somewhat simplify these earlier results. One obtains a simple exact formula for the factorial moments of $\hat{D}_{n,j} = D_{n,j} - 1$ directly from the results of [47]:

$$\mathbb{E}(\hat{D}_{n,j}^s) = s! \frac{\binom{n-j}{s} \left(\frac{s+c_2}{s}\right)}{\binom{j-1+c_2}{s}},$$

with c_1, c_2 as given in (3.8). Hence, by using Stirling's formula for the Gamma function (3.1), for $n \rightarrow \infty$ and $j = j(n) \rightarrow \infty$, the factorial moments of $\hat{D}_{n,j}$ are of mixed Poisson type and Lemma 2 can be applied.

Corollary 3. *The random variable $\hat{D}_{n,j}$, counting the number of descendants minus one of node j in a random increasing tree of size n , has for $n \rightarrow \infty$ and $j = j(n) \rightarrow \infty$, factorial moments of mixed Poisson type with a Gamma mixing distribution $X \stackrel{\mathcal{L}}{=} \gamma(1, 1 + \frac{c_2}{c_1})$, and scale parameter $\lambda_{n,j} = \frac{n-j}{j}$:*

$$\mathbb{E}(\hat{D}_{n,j}^s) = \lambda_{n,j}^s \frac{\Gamma(s+1 + \frac{c_2}{c_1})}{\Gamma(1 + \frac{c_2}{c_1})} (1 + o(1)).$$

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{\hat{D}_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in (0, \infty)$, the random variable $\hat{D}_{n,j}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$, which has a negative binomial distribution.

Remark 7. Note that for fixed n the random variable $\hat{D}_{n,j}/n \xrightarrow{\mathcal{L}} B_j$, where $B_j \stackrel{\mathcal{L}}{=} \beta(1 + \frac{c_2}{c_1}, j-1)$, is asymptotically beta-distributed (see [47]). One readily recovers the mixing distribution X from B_j by taking the limit $j \rightarrow \infty$, using a well known result for beta-distributed random variables:

$$jB_j \rightarrow X, \quad \text{for } j \rightarrow \infty,$$

with convergence of all moments.

Remark 8. Panholzer and Seitz [65] studied labelled families of evolving k -tree models, generalizing simple families of increasing trees. An identical phase change and factorial moments of mixed Poisson type with a Gamma mixing distribution can be observed when studying the number of descendants of specific nodes in labelled families of evolving k -tree models.

The parameter ‘‘descendants of node j ’’ can be modelled also via urn models: we encounter classical Pólya urns with *non-standard initial values*, depending on the number of draws. Note that Mahmoud and Smythe [56] used a similar

approach to study the descendants of node j in recursive trees, for j fixed and $n \rightarrow \infty$.

In order to get an urn model description of the number of descendants of node j one considers the tree evolution process generating random trees of the tree families studied. The probabilities $p(v)$ given in (3.12) can then be translated into the ball replacement matrix of a two-colour urn model. Alternatively, for this task one can use combinatorial descriptions of these tree families, which we will here only figure out for the case d -INCT, i.e., d -ary increasing trees; the other cases can be treated similarly.

Consider such a d -ary increasing tree of size j : there are $1 + (d-1)j$ possible attachment positions (often drawn as external nodes), where a new node can be attached. Exactly d such attachment positions are contained in the subtree rooted j , whereas the other $(d-1)(j-1)$ are not. In the urn model description we will use balls of two colours, black and white. Each white ball will correspond to an attachment position contained in the subtree rooted j , whereas each black ball will correspond to an attachment position that is not contained in the subtree rooted j , which we call here “remaining tree”. This already describes the initial conditions of the urn.

Moreover, during the tree evolution process, when attaching a new node to a position in the subtree of j , then there appear $d-1$ new attachment positions in this subtree, whereas when attaching a new node to a position in the remaining tree, then there appear $d-1$ new attachment positions in the remaining tree. In the urn model description this simply means that when drawing a white ball one adds $d-1$ white balls and when drawing a black ball one adds $d-1$ black balls to the urn. After $n-j$ draws, which correspond to the $n-j$ attachments of nodes in the tree, the number of white balls in the urn is linearly related to the size of the subtree rooted j in a tree of size n . Thus, the following urn model description follows immediately.

Urn III. Consider a Pólya urn with ball replacement matrix

$$M = \begin{pmatrix} \kappa & 0 \\ 0 & \kappa \end{pmatrix}, \quad \kappa = \begin{cases} 1, & \text{RECT,} \\ 1 + \alpha, & \text{GPORT,} \\ d - 1, & d\text{-INCT,} \end{cases}$$

and initial conditions

$$W_0 = \begin{cases} 1, \\ \alpha, \\ d, \end{cases} \quad B_0 = \begin{cases} j - 1, & \text{RECT,} \\ (j - 1)(1 + \alpha), & \text{GPORT,} \\ (j - 1)(d - 1), & d\text{-INCT,} \end{cases}$$

for $1 \leq j \leq n$. The number $D_{n,j}$ of descendants of node j in an increasing tree of size n has the same distribution as the (shifted and scaled) number of white balls W_{n-j} in the Pólya urn after $n-j$ draws,

$$D_{n,j} \stackrel{\mathcal{L}}{=} \begin{cases} W_{n-j}, & \text{RECT,} \\ (W_{n-j} + 1)/\kappa, & \text{GPORT,} \\ (W_{n-j} - 1)/\kappa, & d\text{-INCT.} \end{cases}$$

This implies that the number of white balls in the standard Pólya urn model exhibit a phase transition according to the growth of the initial number of black balls present in the urn compared to the discrete time.

3.4. Node-degrees in plane recursive trees.

Let $X_{n,j}$ denote the random variable counting the outdegree of node j in a generalized plane recursive tree of size n , i.e., a size- n tree from the increasing tree family GPORT defined in Section 3.3.

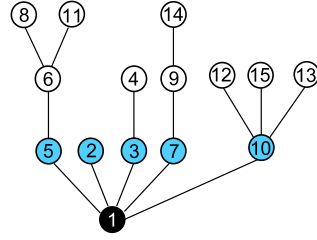


FIG 2. A size 15 plane recursive tree, where the root node $j = 1$ has outdegree five.

It has been shown in [49] using a generating functions approach that the factorial moments of the random variable $X_{n,j}$ are given by

$$\mathbb{E}(X_{n,j}^s) = \frac{\Gamma(s + \alpha)}{\Gamma(\alpha)} \sum_{k=0}^s \binom{s}{k} (-1)^k \frac{\Gamma(n + \frac{s-1-k}{1+\alpha}) \Gamma(j - \frac{1}{1+\alpha})}{\Gamma(j + \frac{s-1-k}{1+\alpha}) \Gamma(n - \frac{1}{1+\alpha})},$$

for $j \geq 2$, with α given in definition (3.11) of the degree-weight generating function $\varphi(t)$. Lemma 2 and an application of Stirling's formula for the Gamma function (3.1) leads to the following result.

Corollary 4. *The random variable $X_{n,j}$, counting the outdegree of node j in a random generalized plane recursive tree of size n , $2 \leq j \leq n$, has for $n \rightarrow \infty$ and $j = j(n) \rightarrow \infty$, falling factorial moments of mixed Poisson type with a Gamma mixing distribution $X \stackrel{\mathcal{L}}{=} \gamma(1, \alpha)$, and scale parameter $\lambda_{n,j} = \left(\frac{n}{j}\right)^{1/(\alpha+1)} - 1$:*

$$\mathbb{E}(X_{n,j}^s) = \lambda_{n,j}^s \frac{\Gamma(s + \alpha)}{\Gamma(\alpha)} (1 + o(1)).$$

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{X_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in (0, \infty)$, the random variable $X_{n,j}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$, which has a negative binomial distribution.

Remark 9. For fixed j , independent of n , a different limit law arises (compare with Remark 5): the random variable $X_{n,j}/n^{1/(\alpha+1)}$ converges for $n \rightarrow \infty$ in distribution to a random variable Z_j characterized by its moments or by its density function; see [49] and also Corollary 5 for details.

The random variable $X_{n,j}$ also allows an urn model description. To get it we consider the tree evolution process for the family GPORT, with $p(v) = \frac{d(v)+\alpha}{(\alpha+1)^{i-1}}$, where $p(v)$ gives the probability that the new node $i+1$ will be attached to node v in a tree of size i , depending on the outdegree $d(v)$ of v . We see that the probability $p(v)$ is proportional to $d(v) + \alpha$. Thus let us think about the quantity $d(v) + \alpha$ as the affinity of node v attracting a new node. The total affinity of all nodes v_k in a tree T of size i is then given by $\sum_{1 \leq k \leq i} (d(v_k) + \alpha) = i - 1 + \alpha i = (\alpha + 1)i - 1$ giving an interpretation of the denominator of $p(v)$. In the two-colour urn model we use white balls describing the affinity of node j to attract new nodes, whereas the black balls describe the affinity of all remaining nodes in the tree to attract new nodes. When considering a tree of size j , it holds $d(j) = 0$ and thus that node j has affinity α , whereas all remaining nodes have total affinity $(\alpha + 1)j - 1 - \alpha = (\alpha + 1)(j - 1)$ to attract a new node. This already yields the initial conditions of the urn.

Each time a new node is attached during the tree evolution process the total affinity of all nodes increases by $\alpha + 1$: α is the affinity of the new node (which has outdegree 0) and by attaching this node the outdegree and thus affinity of one node increases by one. In particular, when a new node is attached to node j the affinity of j increases by one, whereas the total affinity of the remaining nodes increases by α , but if a new node is attached to another node the affinity of j remains unchanged. Thus in the urn model description, when drawing a white ball, one white ball and α black balls are added, and when drawing a black ball, $\alpha + 1$ black balls are added to the urn. The following description is then immediate.

Urn IV. Consider a balanced triangular urn with ball replacement matrix

$$M = \begin{pmatrix} 1 & \alpha \\ 0 & 1 + \alpha \end{pmatrix}, \quad W_0 = \alpha, \quad B_0 = (\alpha + 1)(j - 1),$$

for $1 \leq j \leq n$. The outdegree $X_{n,j}$ of node j in a generalized plane recursive tree of size n has the same distribution as the shifted number of white balls W_{n-j} in the Pólya urn after $n - j$ draws,

$$X_{n,j} \stackrel{\mathcal{L}}{=} W_{n-j} - \alpha.$$

This implies that the number of white balls in the standard Pólya urn model exhibit several phase transitions according to the growth of the initial number of black balls present in the urn with respect to the total number of draws; this will be discussed in detail in a more general setting in Section 4.

3.5. Branching structures in plane recursive trees

Let $X_{n,j,k}$ denote the random variable, which counts the number of size- k branches (= subtrees of size k) attached to the node labelled j in a random increasing tree of size n . The random variables $X_{n,j,k}$ are thus related to the random variable $X_{n,j}$ studied in Section 3.4, counting the outdegree of node labelled j , via

$$X_{n,j} = \sum_{k=1}^{n-j} X_{n,j,k}.$$

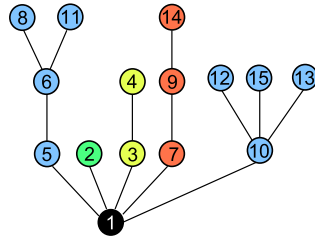


FIG 3. A size 15 plane recursive tree, where the root node has one size-one, one size-two, one size three and two size-four branches.

This parameter was studied in Su et al. [74] for the particular case of the root node $j = 1$ and the instance of random recursive trees: they derived the distribution of $X_{n,1,k}$ and a limit law for it. Furthermore they stated results for joint distributions. The analysis was extended in [48] to increasing tree families generated by a natural growth process (see Section 3.3). In particular, for the family GPOR of generalized plane recursive trees with parameter α given in (3.11), the following result for the factorial moments of $X_{n,j,k}$ was obtained:

$$\mathbb{E}(X_{n,j,k}^s) = \left(\frac{\binom{k-\frac{1}{\alpha+1}}{k-1}}{(\alpha+1)k} \right)^s \cdot \frac{\Gamma(s+\alpha)}{\Gamma(\alpha)} \frac{\binom{j-1-\frac{1}{\alpha+1}}{j-1} \binom{n-ks-1+\frac{s-1}{\alpha+1}}{n-j-ks}}{\binom{n-1}{j-1} \binom{n-1-\frac{1}{\alpha+1}}{n-1}}.$$

In [48] only the case of fixed k was considered. We can easily use Lemma 2 and Stirling's formula for the Gamma function (3.1) to extend the studies given there and to obtain the following result.

Corollary 5. *The random variable $X_{n,j,k}$, counting the number of size- k branches attached to node j in a random generalized plane recursive tree of size n , has for j fixed, $n \rightarrow \infty$ and $1 \leq k \leq n-j$, falling factorial moments of mixed Poisson type with mixing distribution Z_j supported on $[0, \infty)$, uniquely defined by its moment sequence $(\mu_s)_{s \in \mathbb{N}} = \left(\frac{\Gamma(s+\alpha)\Gamma(j-\frac{1}{\alpha+1})}{\Gamma(\alpha)\Gamma(j+\frac{s-1}{\alpha+1})} \right)_{s \in \mathbb{N}}$, and scale parameter*

$$\lambda_{n,j,k} = \frac{n^{\frac{1}{\alpha+1}} \binom{k - \frac{1}{\alpha+1}}{k-1}}{(\alpha+1)^k}.$$

$$\mathbb{E}(X_{n,j,k}^s) = \lambda_{n,j,k}^s \cdot \frac{\Gamma(s + \alpha)\Gamma(j - \frac{1}{\alpha+1})}{\Gamma(\alpha)\Gamma(j + \frac{s-1}{\alpha+1})} (1 + o(1)).$$

- (i) for $\lambda_{n,j,k} \rightarrow \infty$, the random variable $\frac{X_{n,j,k}}{\lambda_{n,j,k}}$ converges in distribution, with convergence of all moments, to Z_j .
- (ii) for $\lambda_{n,j,k} \rightarrow \rho \in (0, \infty)$, the random variable $X_{n,j,k}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho Z_j)$.

Remark 10. The results above can be generalized to growing $j = j(n)$, leading to results similar to our earlier findings for the ordinary outdegree $X_{n,j}$. The random variable Z_j is exactly the limit law of $X_{n,j}/n^{1/(\alpha+1)}$ for fixed j as discussed in Remark 9. The density functions $f_j(x)$ of the Z_j , $j \in \mathbb{N}$, are known explicitly, see [49].

We can interpret our findings in terms of an urn model reminiscent to the urn model for block sizes in k -Stirling permutations. To do this we can extend the description given in Section 3.4 leading to Urn IV. Namely, we require a refinement of describing the affinities of the nodes in a tree to attract new nodes during the tree evolution process.

For doing that we use balls of $k+2$ different colours. Balls of colour 0 describe the affinity of node labelled j attracting a new node to become attached at j . Furthermore, balls of colour i , with $1 \leq i \leq k$, describe the affinity of attracting a new node of nodes that are contained in branches of size i attached to node j , whereas balls of colour $k+1$ describe the affinity of attracting a new node of all remaining nodes, i.e., of nodes that are not contained in the subtree rooted j or nodes contained in a branch of size $\geq k+1$ attached to j . Consider a generalized plane recursive tree of size j : node j has affinity α to attract a new node, whereas the total affinity of all remaining nodes is $(\alpha+1)(j-1)$; this characterizes the initial configuration of the urn.

The ball replacement matrix of the urn can be obtained as follows. When drawing a ball of colour 0, i.e., when attaching a new node to j , the node degree of j increases by one and there appears a new branch of size 1 attached to node j , which means that we add one ball of colour 0 and α balls of colour 1. When drawing a ball of colour i , with $1 \leq i \leq k$, i.e., when attaching a new node to a size- i branch attached to j , this branch transforms into a size- $(i+1)$ branch attached to j , which means that we remove $(\alpha+1)i-1$ balls of colour i (which corresponds to the affinity of a size- i branch) and add $(\alpha+1)(i+1)-1$ balls of colour $i+1$ (which thus correspond to the affinity of a size- $(i+1)$ branch). Furthermore, when drawing a ball of colour $k+1$, i.e., when attaching a new node to node not contained in the subtree rooted j or when attaching a new node to a node contained in a branch of size $\geq k+1$ attached to node j , this does neither affect the affinity of node j nor of its branches of sizes $\leq k$, which means

that we add $\alpha + 1$ balls of colour $k + 1$. The following urn model description immediately follows.

Urn V. Consider a balanced urn with balls of $k + 2$ colors and let the random vector $(U_{n,0}, U_{n,1}, \dots, U_{n,k+1})$ count the number of balls of each color at time n with $(k + 2) \times (k + 2)$ ball replacement matrix M given by

$$M = \begin{pmatrix} 1 & \alpha & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & -\alpha & 2(\alpha+1)-1 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & -(2(\alpha+1)-1) & 3(\alpha+1)-1 & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & -((\alpha+1)(k-1)-1) & (\alpha+1)k-1 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 & -((\alpha+1)k-1) & (\alpha+1)(k+1)-1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1+\alpha \end{pmatrix}.$$

The initial configuration of the urn (it is convenient to start here at time 0) is given by $(U_{0,0}, U_{0,1}, \dots, U_{0,k+1}) = (\alpha, 0, \dots, 0, (\alpha + 1)(j - 1))$. The random variables $U_{n,i}$, with $1 \leq i \leq k$, described by the urn model are related to the random variables $X_{n,j,i}$, $1 \leq i \leq k$, which count the number of size- i branches attached to the node labelled j in a random generalized plane recursive tree of size n , as follows:

$$U_{n-j,i} = ((\alpha + 1)i - 1)X_{n,j,i}, \quad 1 \leq i \leq k.$$

Moreover, $U_{n-j,0}$ is related to the outdegree $X_{n,j}$ via $U_{n-j,0} = X_{n,j} + \alpha$.

This implies that the random variables $U_{n,i}$ occurring in the urn model undergo a phase transition according to the growth of k with respect to n , from continuous to discrete.

3.6. Distribution of table sizes in the Chinese restaurant process

The Chinese restaurant process with parameters a and θ is a discrete-time stochastic process, whose value at any positive integer time n is one of the B_n partitions of the set $[n] = \{1, 2, 3, \dots, n\}$ (see Pitman [69]). The parameters a and θ satisfy $0 < a < 1$ and $\theta > -a$. Here B_n denotes the Bell number counting the number of partitions of an n -element set $B_0 = B_1 = 1, B_2 = 2, B_3 = 5$, etc.² One imagines a Chinese restaurant with an infinite number of tables, and each table has an infinite number of seats. In the beginning the first customer takes place at the first table. At each discrete time step a new customer arrives and either joins one of the existing tables, or he takes place at the next empty table in line. Each table corresponds to a block of a random partition. In the beginning at time $n = 1$, the trivial partition $\{\{1\}\}$ is obtained with probability 1. Given a partition T of $[n]$ with $|T| = k$ parts $t_i, 1 \leq i \leq k \leq n$, of sizes $|t_i|$. At

²See sequence [A000110](#) in OEIS.

time $n + 1$ the element $n + 1$ is either added to one of the existing parts $t_i \in T$ with probability

$$\mathbb{P}\{n + 1 <_c t_i\} = \frac{|t_i| - a}{n + \theta}, \quad 1 \leq i \leq k,$$

or added to the partition T as a new singleton block with probability

$$\mathbb{P}\{n + 1 <_c t_{|T|+1}\} = \frac{\theta + k \cdot a}{n + \theta}.$$

This model thus assigns a probability to any particular partition T of $[n]$. We are interested in the distribution of the random variable $C_{n,j}$, counting the number of parts of size j in a partition of $[n]$ generated by the Chinese restaurant process.

We will not directly study the Chinese restaurant process, but in order to analyse the number of tables of a certain size we study instead a variant of the growth rule for generalized plane recursive trees as introduced in Section 3.3. Combinatorially, we consider a family $\mathcal{T}_{\alpha,\beta}$ of generalized plane recursive trees, where the degree-weight generating function $\vartheta(t) = \frac{1}{(1-t)^\beta}$, $\beta > 0$, associated to the root of the tree, is different to the one for non-root nodes in the tree, $\varphi(t) = \frac{1}{(1-t)^\alpha}$, $\alpha > 0$. Then, the family $\mathcal{T}_{\alpha,\beta}$ is closely related to the corresponding family \mathcal{T} of generalized plane recursive trees with degree-weight generating $\varphi(t) = \frac{1}{(1-t)^\alpha}$, $\alpha > 0$, via the following formal recursive equations:

$$\mathcal{T}_{\alpha,\beta} = \mathbb{1} \times \vartheta(\mathcal{T}), \quad \mathcal{T} = \mathbb{1} \times \varphi(\mathcal{T}). \quad (3.13)$$

The weight $w(T)$ of a tree $T \in \mathcal{T}_{\alpha,\beta}$ is then defined as

$$w(T) := \vartheta_{d(\text{root})} \prod_{v \in T \setminus \{\text{root}\}} \varphi_{d(v)},$$

where $d(v)$ denotes the outdegree of node v . Thus, the generating functions $T_{\alpha,\beta}(z) = \sum_{n \geq 1} T_{\alpha,\beta;n} \frac{z^n}{n!}$ and $T(z) = \sum_{n \geq 1} T_n \frac{z^n}{n!}$ of the total weight of size- n trees in $\mathcal{T}_{\alpha,\beta}$ and \mathcal{T} , respectively, satisfies the differential equations

$$T'_{\alpha,\beta}(z) = \vartheta(T(z)), \quad T'(z) = \varphi(T(z)).$$

Moreover, the tree evolution process to generate a random tree of arbitrary given size in the family \mathcal{T} described in Section 3.3 can be extended in the following way to generate a random tree in the family $\mathcal{T}_{\alpha,\beta}$. The process, evolving in discrete time, starts with the root labelled by zero. At step $n + 1$, with $n \geq 0$, the node with label $n + 1$ is attached to any previous node v with outdegree $d(v)$ of the already grown tree with probabilities $p(v)$, which are given for non-root nodes v by

$$\mathbb{P}\{n + 1 <_c v\} = \frac{d(v) + \alpha}{\beta + (\alpha + 1)n},$$

and for the root by

$$\mathbb{P}\{n + 1 <_c \text{root}\} = \frac{d(\text{root}) + \beta}{\beta + (\alpha + 1)n}.$$

This growth process is similar to the Chinese restaurant process considered before. Indeed, if we remove the root labelled zero, the remaining branches contain the nodes with labels given by $[n] = \{1, \dots, n\}$.

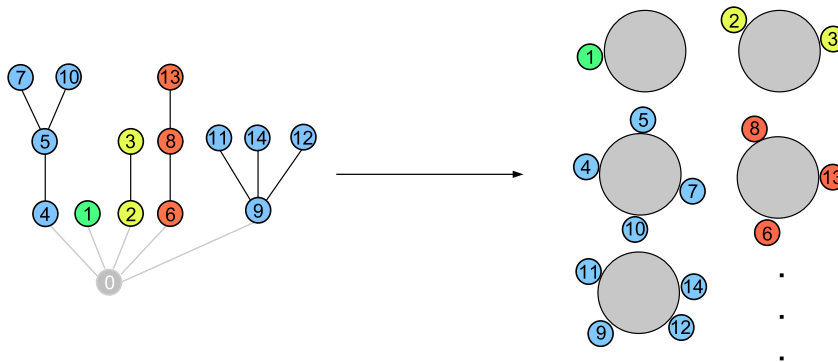


FIG 4. A size 15 plane recursive tree, where the root node labelled zero has one size-one, one size-two, one size three, two size-four branches and the corresponding table structure in the Chinese restaurant model.

Proposition 3 (Chinese restaurant process and generalized plane recursive trees). *A random partition of $\{1, \dots, n\}$ generated by the Chinese restaurant process with parameters a and θ can be generated equivalently by the growth process of the family of generalized plane recursive trees $\mathcal{T}_{\alpha, \beta}$ when generating such a tree of size $n + 1$. The parameters a, θ and $\alpha, \beta > 0$, respectively, are related via*

$$a = \frac{1}{1 + \alpha}, \quad \theta = \frac{\beta}{1 + \alpha}.$$

Remark 11. In above relation, θ cannot be negative, since β is assumed to be positive. The above correspondence can be extended to the full range $\beta > -1$ using a different degree-weight generating function $\vartheta(t)$ for the root node. Assume that $-1 < \beta \leq 0$. Then, we cannot directly use $\vartheta(t) = (1 - t)^{-\beta} = 1 + \beta t + \dots$ due to the negative or zero weight. Since the root connectivity is similar to the choice $\beta \mapsto 1 + \beta$ for an outdegree of the root larger than one, we use a shifted connectivity of the root node:

$$\vartheta(t) = 1 + \int_0^t \frac{1}{(1 - x)^{1+\beta}} dx = 1 + \frac{1}{\beta} \left(\frac{1}{(1 - t)^\beta} - 1 \right) = 1 + \sum_{k \geq 1} \binom{\beta + k}{k - 1} \frac{t^k}{k},$$

for $-1 < \beta < 0$, and

$$\vartheta(t) = 1 - \log(1 - t) = 1 + \sum_{k \geq 1} \frac{t^k}{k},$$

for $\beta = 0$.

Proof of Proposition 3. Assume that a size- n tree T of the family $\mathcal{T}_{\alpha,\beta}$ with labels $\{0, 1, \dots, n\}$ has k branches t_i , $1 \leq i \leq k$, of sizes $|t_i|$. By the considerations of [63], at time $n + 1$ the element $n + 1$ is either attached to one of the existing non-root nodes v with probability

$$\mathbb{P}\{n + 1 <_c v\} = \frac{d(v) + \alpha}{\beta + (\alpha + 1)n},$$

or to the root of the tree with probability

$$\mathbb{P}\{n + 1 <_c \text{root}\} = \frac{d(\text{root}) + \beta}{\beta + (\alpha + 1)n} = \frac{k + \beta}{\beta + (\alpha + 1)n} = \frac{\frac{\beta}{\alpha+1} + k \cdot \frac{1}{\alpha+1}}{n + \frac{\beta}{\alpha+1}}.$$

Consequently, element $n + 1$ is attached to one of the branches $t_i \in T$ with probability

$$\begin{aligned} \mathbb{P}\{n + 1 <_c t_i\} &= \sum_{v \in t_i} \mathbb{P}\{n + 1 <_c v\} = \sum_{v \in t_i} \frac{d(v) + \alpha}{\beta + (\alpha + 1)n} \\ &= \frac{|t_i| - 1 + |t_i|\alpha}{\beta + (\alpha + 1)n} = \frac{|t_i| - \frac{1}{\alpha+1}}{n + \frac{\beta}{\alpha+1}}, \end{aligned}$$

Thus, setting $a = \frac{1}{1+\alpha}$ and $\theta = \frac{\beta}{1+\alpha}$ proves the stated result. \square

Theorem 3.1. *The random variable $C_{n,j}$, counting the number of parts of size j in a partition of $\{1, \dots, n\}$ generated by the Chinese restaurant process with parameters a and θ , is distributed as $X_{n+1,j}$, which counts the number of branches of size j attached to the root of a random size- $(n+1)$ generalized plane recursive tree of the family $\mathcal{T}_{\alpha,\beta}$, with $a = \frac{1}{1+\alpha}$ and $\theta = \frac{\beta}{1+\alpha}$:*

$$C_{n,j} \stackrel{\mathcal{L}}{=} X_{n+1,j}.$$

Assume that $\beta > 0$. Then, $X_{n+1,j}$ has falling factorial moments of mixed Poisson type with mixing distribution Z supported on $[0, \infty)$, and scale parameter

$$\lambda_{n,j} = \frac{n^{\frac{1}{\alpha+1}} \binom{j-1-\frac{1}{\alpha+1}}{j-1}}{(\alpha+1)^j}:$$

$$\mathbb{E}(X_{n+1,j}^s) = \lambda_{n,j}^s \cdot \frac{\Gamma(s + \beta) \Gamma(\frac{\beta}{\alpha+1})}{\Gamma(\beta) \Gamma(\frac{\beta+s}{\alpha+1})} (1 + o(1)).$$

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{X_{n+1,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to Z .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in (0, \infty)$, the random variable $X_{n+1,j}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho Z)$.

Remark 12. A similar result holds true for $-1 < \beta \leq 0$. The analysis is identical, but one has to use the adapted degree-weight generating functions stated in Remark 11.

Proof of Theorem 3.1. We can study $X_{n+1,j}$, which counts the number of branches of size j attached to the root of a size- $(n+1)$ tree using the variable v as a marker and the generating function

$$T_{\alpha,\beta}(z, v) = \sum_{n \geq 1} T_{\alpha,\beta;n} \mathbb{E}(v^{X_{n,j}}) \frac{z^n}{n!}.$$

We have

$$T'_{\alpha,\beta}(z, v) = \vartheta(T(z) - \frac{T_j}{j!} z^j (1-v)), \quad T'(z) = \varphi(T(z)).$$

Solving the differential equation for $T(z)$ leads to

$$T'_{\alpha,\beta}(z, v) = \frac{1}{(1 - T(z) + \frac{T_j}{j!} z^j (1-v))^\beta}, \quad T(z) = 1 - (1 - (\alpha+1)z)^{\frac{1}{\alpha+1}}.$$

We can access the s^{th} moment of $X_{n+1,j}$ as follows:

$$\mathbb{E}(X_{n+1,j}^s) = s! \frac{[z^n w^s] T'_{\alpha,\beta}(z, v)}{\frac{T_{\alpha,\beta;n}}{n!}},$$

where we set $w = v - 1$. Consequently, an expansion of $T'_{\alpha,\beta}(z, v)$ at $w = 0$ gives

$$\mathbb{E}(X_{n+1,j}^s) = \frac{n!s!}{T_{\alpha,\beta;n}} \cdot \left(\frac{T_j}{j!}\right)^s \binom{\beta-1+s}{s} [z^{n-j s}] \frac{1}{(1-T(z))^{\beta+s}}.$$

Since $\frac{T_{\alpha,\beta;n}}{n!} = (\alpha+1)^n \binom{\beta}{\alpha+1}_n^{+n-1}$, we obtain the explicit result

$$\mathbb{E}(X_{n+1,j}^s) = \left(\frac{\binom{j-1-\frac{1}{\alpha+1}}{j-1}}{(\alpha+1)j}\right)^s \cdot \frac{s!}{\binom{\beta}{\alpha+1}_n^{+n-1}} \cdot \binom{\beta-1+s}{s} \binom{n-j s - 1 + \frac{\beta+s}{\alpha+1}}{n-j s}.$$

An application of Stirling's formula for the Gamma function yields then the stated result. \square

4. Triangular urn models

During the study of node-degree in generalized plane recursive trees in Section 3.4 we encountered a triangular urn model leading to factorial moments of mixed Poisson type. Here we study a more general triangular urn.

Urn VI. Consider a balanced triangular urn model with ball replacement matrix

$$M = \begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix}, \quad \alpha, \beta, \gamma \in \mathbb{N}, \quad \gamma = \alpha + \beta \in \mathbb{N}.$$

The initial configuration of the urn consists of w_0 white balls and b_0 black balls, and the random variable W_n counts the number of white balls in the urn after n draws.

This urn model has been studied by Puyhaubert [21, 72] who derived the probability mass function of W_n and a limit law for $n \rightarrow \infty$. The results of [21, 72] were extended by Janson [37] to unbalanced triangular urn models. Here, using a simple closed formula for the rising factorial moments of W_n , we point out several phase transitions, involving amongst others moments of mixed Poisson type, for non-standard initial values $b_0 = b_0(n)$, which may depend on the discrete time n . Due to the balanced nature of the urn the total number T_n of balls after n draws is a deterministic quantity:

$$T_n = T_0 + n \cdot \gamma, \quad n \geq 0, \quad T_0 = w_0 + b_0.$$

Our starting point is the analysis of the normalized number of white balls $X_n = W_n/\alpha$, such that $X_0 = w_0/\alpha$. Let \mathcal{F}_n denote the σ -field generated by the first n steps. Moreover, denote by $\Delta_n = X_n - X_{n-1} \in \{0, 1\}$ the increment at step n . We have

$$\mathbb{E}(X_n | \mathcal{F}_{n-1}) = \mathbb{E}(X_{n-1} + \Delta_n | \mathcal{F}_{n-1}) = X_{n-1} + \mathbb{E}(\Delta_n | \mathcal{F}_{n-1}).$$

Since the probability that a new white ball is generated at step n is proportional to the number $W_{n-1} = X_{n-1} \cdot \alpha$ of existing white balls (at step $n-1$), we obtain further

$$\mathbb{E}(X_n | \mathcal{F}_{n-1}) = X_{n-1} + \frac{X_{n-1} \cdot \alpha}{T_{n-1}} = \frac{T_{n-1} + \alpha}{T_{n-1}} W_{n-1}, \quad n \geq 1.$$

Hence, let

$$\mathcal{X}_n = X_n \cdot \prod_{k=0}^{n-1} \frac{T_k}{T_k + \alpha} = X_n \cdot \frac{\binom{n-1 + \frac{T_0}{\gamma}}{n}}{\binom{n-1 + \frac{T_0 + \alpha}{\gamma}}{n}}.$$

Then

$$\mathbb{E}(\mathcal{X}_n | \mathcal{F}_{n-1}) = \frac{\binom{n-1 + \frac{T_0}{\gamma}}{n}}{\binom{n-1 + \frac{T_0 + \alpha}{\gamma}}{n}} \cdot \frac{T_{n-1} + \alpha}{T_{n-1}} W_{n-1} = \mathcal{X}_{n-1}, \quad n \geq 1.$$

Consequently, \mathcal{X}_n is a positive martingale. By taking the unconditional expectation, this implies that the expected value of X_n is given by

$$\mathbb{E}(X_n) = \frac{\binom{n-1 + \frac{T_0 + \alpha}{\gamma}}{n}}{\binom{n-1 + \frac{T_0}{\gamma}}{n}} \cdot \mathbb{E}(X_0) = \frac{\binom{n-1 + \frac{T_0 + \alpha}{\gamma}}{n}}{\binom{n-1 + \frac{T_0}{\gamma}}{n}} \cdot X_0.$$

More generally, we similarly have for any positive integer s

$$\begin{aligned}\mathbb{E}\left(\binom{X_n + s - 1}{s} \middle| \mathcal{F}_{n-1}\right) &= \binom{X_{n-1} + s - 1}{s} + \binom{X_{n-1} + s - 1}{s-1} \frac{\alpha X_{n-1}}{T_{n-1}} \\ &= \binom{X_{n-1} + s - 1}{s} \frac{T_{n-1} + s\alpha}{T_{n-1}}.\end{aligned}$$

Hence, this implies that the s^{th} binomial moment is given by

$$\begin{aligned}\mathbb{E}\left(\binom{X_n + s - 1}{s}\right) &= \frac{\binom{n-1+\frac{T_0+s\alpha}{\gamma}}{n}}{\binom{n-1+\frac{T_0}{\gamma}}{n}} \cdot \binom{X_0 + s - 1}{s} \\ &= \frac{\Gamma(n + \frac{w_0+b_0+s\alpha}{\gamma})\Gamma(\frac{w_0+b_0}{\gamma})\Gamma(\frac{w_0}{\alpha} + s)}{\Gamma(\frac{w_0+b_0+s\alpha}{\gamma})\Gamma(n + \frac{w_0+b_0}{\gamma})\Gamma(\frac{w_0}{\alpha})\Gamma(s+1)}.\end{aligned}$$

Theorem 4.1. *The s^{th} rising factorial moment of the random variable $X_n = W_n/\alpha$, where W_n counts the number of white ball in a balanced triangular urn with ball replacement matrix given by $\begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix}$, $\alpha, \beta, \gamma \in \mathbb{N}$, $\gamma = \alpha + \beta$, is given by the exact formula*

$$\mathbb{E}(X_n^{\overline{s}}) = \frac{\Gamma(n + \frac{w_0+b_0+s\alpha}{\gamma})\Gamma(\frac{w_0+b_0}{\gamma})\Gamma(\frac{w_0}{\alpha} + s)}{\Gamma(\frac{w_0+b_0+s\alpha}{\gamma})\Gamma(n + \frac{w_0+b_0}{\gamma})\Gamma(\frac{w_0}{\alpha})},$$

where w_0, b_0 denote the initial number of white and black balls, respectively. The factorial moments of $\hat{X}_n = X_n - \frac{w_0}{\alpha}$ are for $\min\{n, b_0\} \rightarrow \infty$ asymptotically of mixed Poisson type with a gamma mixing distribution $X \stackrel{L}{=} \gamma(\frac{w_0}{\alpha}, 1)$, and scale parameter $\lambda_{n,b_0} = \left(\frac{n+\frac{b_0}{\gamma}}{\frac{b_0}{\gamma}}\right)^{\frac{\alpha}{\gamma}} - 1$,

$$\mathbb{E}(\hat{X}_n^{\overline{s}}) = (\lambda_{n,b_0})^s \cdot \frac{\Gamma(\frac{w_0}{\alpha} + s)}{\Gamma(\frac{w_0}{\alpha})} (1 + o(1)).$$

(i) for $\lambda_{n,b_0} \rightarrow \infty$, the random variable $\frac{\hat{X}_n}{\lambda_{n,b_0}}$ converges in distribution, with convergence of all moments, to X .

(ii) for $\lambda_{n,b_0} \rightarrow \rho \in (0, \infty)$, the random variable \hat{X}_n converges in distribution, with convergence of all moments, to $Y \stackrel{L}{=} MPo(\rho X)$.

Remark 13. It is well known from the works of Puyhaubert [21, 72] and Janson [37], that for fixed b_0 the random variable $X_n/n^{\frac{\alpha}{\gamma}}$ tends to a random variable Z , depending on the initial condition w_0, b_0 and also α and γ . Its power moments are given by

$$\mathbb{E}(Z^s) = \frac{\Gamma(\frac{w_0+b_0}{\gamma})\Gamma(\frac{w_0}{\alpha} + s)}{\Gamma(\frac{w_0+b_0+s\alpha}{\gamma})\Gamma(\frac{w_0}{\alpha})}, \quad s \geq 1;$$

for more details about the nature of this random variable we refer the reader to [21, 37]. This result can easily be re-obtained using the explicit expression

for the rising factorial moments of X_n and the method of moments. We obtain the gamma mixing distribution $X \stackrel{\mathcal{L}}{=} \gamma(\frac{w_0}{\alpha}, 1)$ from $Z = Z_{b_0}$ as follows:

$$b_0^{\frac{\alpha}{\gamma}} Z_{b_0} \xrightarrow{\mathcal{L}} X, \quad \text{for } b_0 \rightarrow \infty,$$

with convergence of all moments.

Proof of Theorem 4.1. Let Y denote a random variable with *rising factorial moments* $\mathbb{E}(Y^{\bar{s}}) = \mathbb{E}(Y(Y+1)\dots(Y+s-1))$ satisfying an expansion of mixed Poisson type, $\mathbb{E}(Y^{\bar{s}}) = \rho^s \cdot \mu_s$, for $s \geq 1$, with $\mu_s \geq 0$. We obtain the (falling) factorial moments using the binomial theorem for rising factorials (see [30]):

$$\begin{aligned} x^s &= (x-s+1)^{\bar{s}} = \sum_{\ell=1}^s \binom{s}{\ell} \cdot x^{\bar{\ell}} \cdot (-s+1)^{\overline{s-\ell}} \\ &= \sum_{\ell=1}^s \binom{s}{\ell} x^{\bar{\ell}} (-1)^{s-\ell} (s-1)^{\overline{s-\ell}}, \quad s \geq 1. \end{aligned}$$

Moreover, we can obtain the rising factorial moments of the shifted random variable $\hat{X}_n = X_n - \frac{w_0}{\alpha}$ by using again the binomial theorem

$$(x+c)^{\bar{\ell}} = \sum_{j=0}^{\ell} \binom{\ell}{j} \cdot x^{\bar{j}} \cdot c^{\overline{\ell-j}}, \quad \ell \geq 0.$$

This implies that we can express the factorial moments of \hat{X}_n in terms of the rising factorial moments of X_n by combining the two identities above in the following way.

$$\begin{aligned} \mathbb{E}(\hat{X}_n^{\bar{s}}) &= \sum_{\ell=1}^s \binom{s}{\ell} \mathbb{E}(\hat{X}_n^{\bar{\ell}}) (-1)^{s-\ell} (s-1)^{\overline{s-\ell}} \\ &= \sum_{\ell=1}^s \binom{s}{\ell} (-1)^{s-\ell} (s-1)^{\overline{s-\ell}} \sum_{j=0}^{\ell} \binom{\ell}{j} \mathbb{E}(X_n^{\bar{j}}) (-w_0/\alpha)^{\overline{\ell-j}}. \end{aligned}$$

Next we use the asymptotic expansion of the rising factorial moments of X_n ,

$$\mathbb{E}(X_n^{\bar{s}}) = (\lambda_{n,b_0} + 1)^s \frac{\Gamma(\frac{w_0}{\alpha} + s)}{\Gamma(\frac{w_0}{\alpha})} (1 + o(1)), \quad s \geq 1,$$

where $\lambda_{n,b_0} = \left(\frac{n+\frac{b_0}{\gamma}}{\frac{b_0}{\gamma}}\right)^{\frac{\alpha}{\gamma}} - 1$. Interchanging summations, and collecting powers of λ_{n,b_0} leads to the expansion

$$\begin{aligned} \mathbb{E}(\hat{X}_n^{\bar{s}}) &= s! \sum_{i=0}^s \lambda_{n,b_0}^i \left[\sum_{j=i}^s \binom{j}{i} (-1)^j \binom{\frac{w_0}{\alpha} + j - 1}{j} \times \right. \\ &\quad \left. \times \sum_{\ell=j}^s \binom{\ell-1}{\ell-j} \binom{-\frac{w_0}{\alpha}}{s-\ell} (-1)^\ell \right] (1 + o(1)). \end{aligned}$$

Next, using the hypergeometric form of the Vandermonde convolution, see [30, p. 212], we obtain for the inner sum

$$\sum_{\ell=j}^s \binom{\ell-1}{\ell-j} \binom{-\frac{w_0}{\alpha}}{s-\ell} (-1)^\ell = \frac{\binom{-\frac{w_0}{\alpha}}{s-j} \binom{j-s-\frac{w_0}{\alpha}}{j}}{\binom{-\frac{w_0}{\alpha}+j-1}{j}}.$$

We get further

$$\begin{aligned} s! \sum_{j=i}^s \binom{j}{i} \binom{-\frac{w_0}{\alpha}}{s-j} \binom{j-s-\frac{w_0}{\alpha}}{j} (-1)^j &= \delta_{s,i} \cdot s! (-1)^s \binom{-\frac{w_0}{\alpha}}{s} \\ &= \delta_{s,i} \cdot \frac{\Gamma(\frac{w_0}{\alpha} + s)}{\Gamma(\frac{w_0}{\alpha})}, \end{aligned}$$

where $\delta_{s,i}$ denotes the Kronecker-delta function. This proves the stated result. \square

5. Mixed Poisson-Rayleigh laws

In the analysis of various combinatorial objects as, e.g., lattice paths, trees and mappings, the Rayleigh distribution occurs frequently. In this section we give several examples, where during the study of such objects, a mixed Poisson distribution with Rayleigh mixing distribution occurs in a natural way. Apart from the first example, the occurrence and proof of the mixed Poisson distribution is novel, best to our knowledge.

5.1. The number of inversions in labelled tree families

Consider a rooted labelled tree T , where the nodes of T are labelled with distinct integers (usually of the set $\{1, 2, \dots, |T|\}$, with $|T|$ the size, i.e., the number of vertices, of T). An *inversion* in T is a pair (i, j) of vertices (we may always identify a vertex with its label), such that $i > j$ and i lies on the unique path from the root node of T to j (thus i is an ascendant of j or, equivalently, j is a descendant of i). Given a tree family, we introduce the r.v. $I_{n,j}$, which counts, for a random tree of size n , the number of inversions induced by the node labelled j , $1 \leq j \leq n$, i.e., it counts the number of inversions of the kind (i, j) , with $i > j$ an ancestor of j . See Figure 5 for an illustration of the quantity considered.

Panholzer and Seitz [64] studied the r.v. $I_{n,j}$ for random trees of so-called labelled simply generated tree families (see, e.g., [24]; note that in the probabilistic literature such tree models are more commonly called Galton-Watson trees), which contain many important tree families as, e.g., ordered, unordered, binary and cyclic labelled trees as special instances.

Formally, a class \mathcal{T} of labelled simply generated trees is defined in the following way: One chooses a sequence $(\varphi_\ell)_{\ell \geq 0}$ (the so-called *degree-weight sequence*)

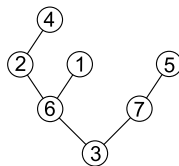


FIG 5. A binary labelled tree of size 7 with a total number of 6 inversions, namely (3, 1), (6, 1), (3, 2), (6, 2), (6, 4), (7, 5). Thus two inversions each are induced by the nodes 1 and 2, whereas one inversion each is induced by the nodes 4 and 5.

of nonnegative real numbers with $\varphi_0 > 0$. Using this sequence, the *weight* $w(T)$ of each labelled ordered tree (i.e., each labelled rooted tree, in which the children of each node are ordered from left to right) is defined by $w(T) := \prod_{v \in T} \varphi_{d(v)}$, where by $v \in T$ we mean that v is a vertex of T and $d(v)$ denotes the number of children of v (i.e., the outdegree of v). The family \mathcal{T} associated to the degree-weight sequence $(\varphi_\ell)_{\ell \geq 0}$ then consists of all trees T (or all trees T with $w(T) \neq 0$) together with their weights. We let $T_n := \sum_{|T|=n} w(T)$ denote the *total weight* of all trees of size n in \mathcal{T} ; for many important simply generated tree families, $(T_n)_{n \geq 1}$ is a sequence of natural numbers, and then the total weight T_n can be interpreted simply as the number of trees of size n in \mathcal{T} .

When analysing parameters in a simply generated tree family \mathcal{T} it is common to assume the *random tree model for weighted trees*, i.e., when speaking about a random tree of size n one assumes that each tree T in \mathcal{T} of size n is chosen with a probability proportional to its weight $w(T)$, i.e., is chosen with probability $\frac{w(T)}{T_n}$. Under mild conditions on the degree-weight sequence $(\varphi_\ell)_{\ell \geq 0}$ of a family \mathcal{T} of labelled simply generated trees and assuming the random tree model, in [64] the following asymptotic formula for the factorial moments of $I_{n,j}$ has been obtained:

$$\mathbb{E}(I_{n,j}^s) = \Gamma\left(\frac{s}{2} + 1\right) \left(\frac{2}{\kappa}\right)^{\frac{s}{2}} \frac{(n-j)^s}{n^{\frac{s}{2}}} \cdot (1 + o(1)),$$

where the constant κ depends on the particular tree family, i.e., on the degree-weight sequence, and is given in [64]. Consequently, an application of Lemma 2 and taking into account Example 3 yields the following result, which adds to the results of [64] the characterization of the limiting distribution as a mixed Poisson distribution.

Corollary 6. *The random variable $I_{n,j}$, which counts the number of inversions induced by node j in a random labelled simply generated tree of size n has for $n \rightarrow \infty$ and arbitrary $1 \leq j = j(n) \leq n$ asymptotically factorial moments of mixed Poisson type with a Rayleigh mixing distribution X and scale parameter $\lambda_{n,j} = \sqrt{\frac{1}{\kappa} \frac{n-j}{\sqrt{n}}}$ (with constant κ given in [64]),*

$$\mathbb{E}(I_{n,j}^s) = \lambda_{n,j}^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right) \cdot (1 + o(1)).$$

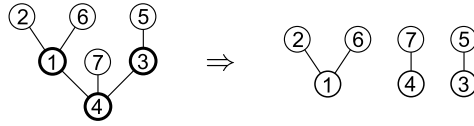


FIG 6. A labelled tree of size 7 with three min-records: nodes 1, 3, and 4. This decomposes the original tree into three record subtrees, one of them is of size 3 and two of them are of size 2.

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{I_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in (0, \infty)$, the random variable $I_{n,j}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$.

Moreover, the random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ converges for $\rho \rightarrow \infty$, after scaling, to its mixing distribution $X: \frac{Y}{\rho} \stackrel{\mathcal{L}}{\rightarrow} X$, with convergence of all moments.

Remark 14. Considering $\lambda_{n,j}$, the critical phase occurs at $j = n - \Theta(\sqrt{n})$: $\mathbb{E}(I_{n,j}) \rightarrow \infty$, for $n - j \gg \sqrt{n}$, whereas $\mathbb{E}(I_{n,j}) \rightarrow 0$, for $n - j \ll \sqrt{n}$.

5.2. Record-subtrees in Cayley trees

Given a rooted labelled tree T , a *min-record* (or simply *record*, for short) is a node $x \in T$, which has the smallest label amongst all nodes on the (unique) path from the root-node of T to x . Let us assume that $\{r_1, \dots, r_k\}$ is the set of records of T ; then this set naturally induces a decomposition of the tree T into what is called here *record-subtrees* $\{S_1, \dots, S_k\}$: S_i , $1 \leq i \leq k$, is the largest subtree rooted at the record r_i not containing any of the remaining records $r_1, \dots, r_{i-1}, r_{i+1}, \dots, r_k$. In other words, a record-subtree S is a maximal subtree (i.e., it is not properly contained in another such subtree) of T with the property that the root-node of S has the smallest label amongst all nodes of S . See Figure 6 for an illustration of these quantities.

In the following we will study the occurrence of record-subtrees of a given size for one of the most natural random tree models, namely random rooted labelled unordered trees, often called random Cayley trees. A Cayley tree is a rooted tree T , where the nodes of T are labelled with distinct integers of $\{1, 2, \dots, |T|\}$ and where the children of any node $x \in T$ are not equipped with any left-to-right ordering (i.e., we may think that each node in T has a possibly empty set of children). Combinatorially, the family \mathcal{T} of Cayley trees can be described formally via the SET construction as

$$\mathcal{T} = \text{O} * \text{SET}(\mathcal{T}). \tag{5.1}$$

Note that Cayley trees are a particular family of labelled simply generated trees as described in Section 5.1, where the degree-weight sequence $(\varphi_\ell)_{\ell \geq 0}$ is given

by $\varphi_\ell = \frac{1}{\ell}$. It is well-known that there are exactly $T_n = n^{n-1}$ different Cayley trees of size n (see, e.g., [24, 73]) and in the random tree model, which we will always assume here, each of these trees may occur with the same probability when considering a size- n tree.

The r.v. R_n , counting the number of records in a random size- n Galton-Watson tree (i.e., in a simply generated tree), has been studied by Janson [36] showing (after a suitable scaling by $\frac{1}{\sqrt{n}}$) a Rayleigh limiting distribution result; in particular, for Cayley trees it holds $\frac{R_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \text{Rayleigh}(1)$. Here we introduce the r.v. $R_{n,j}$, which counts the number of record-subtrees of size j in a random Cayley tree of size n . Of course, the random variables $R_{n,j}$, $1 \leq j \leq n$, are a refinement of R_n and are related by the identity

$$R_n = \sum_{j=1}^n R_{n,j}.$$

As has been pointed out already in [36], records in trees are closely related to a certain node removal procedure for trees. Starting with a tree T one chooses a node $x \in T$ at random and cuts off the subtree T'' of T rooted at x , and iterates this cutting procedure with the remaining subtree T' until only the empty subtree remains. The r.v. $C_n^{[v]}$ counting the number of (vertex) cuts required to cut-down the whole tree by this cutting procedure when starting with a random Cayley tree of size n is then distributed as R_n , i.e., $R_n \stackrel{\mathcal{L}}{=} C_n^{[v]}$. We can extend this relation by considering the r.v. $C_{n,j}^{[v]}$ counting the number of subtrees of size j , which are cut-off during the (vertex) cutting procedure when starting with a random Cayley tree of size n , where it holds $R_{n,j} \stackrel{\mathcal{L}}{=} C_{n,j}^{[v]}$. This can be seen easily by means of coupling arguments given in [36]: consider the node-removal procedure, where, starting with a tree T , in each step the node with smallest label amongst all nodes in the remaining tree is selected and together with all its descendants detached from the tree. Then it holds that node x is a min-record in the tree T if and only if node x is selected as a vertex cut during this node-removal procedure and in this case the record-subtree (and thus its size) rooted at x corresponds to the subtree (with its respective size), which is removed in this cut.

We will show that $R_{n,j}$ and thus also $C_{n,j}^{[v]}$ has factorial moments of mixed Poisson type yielding the following theorem.

Theorem 5.1. *The random variable $R_{n,j}$, counting the number of record-subtrees of size j in a random Cayley tree of size n , has, for $n \rightarrow \infty$ and arbitrary $1 \leq j = j(n) \leq n$, asymptotically factorial moments of mixed Poisson type with a Rayleigh mixing distribution X and scale parameter $\lambda_{n,j} = \frac{\sqrt{n}j^{j-1}}{j!e^j}$:*

$$\mathbb{E}(R_{n,j}^s) = \lambda_{n,j}^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right) \cdot (1 + o(1)), \quad \text{for } j = o(n), \quad \mathbb{E}(R_{n,j}^s) = \mathcal{O}(\lambda_{n,j}^s).$$

(i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{R_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .

(ii) for $\lambda_{n,j} \rightarrow \rho \in [0, \infty)$, the random variable $R_{n,j}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$.

Moreover, the random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ converges, for $\rho \rightarrow \infty$, after scaling, to its mixing distribution $X: \frac{Y}{\rho} \stackrel{\mathcal{L}}{\rightarrow} X$, with convergence of all moments.

Remark 15. Stirling's formula for the Gamma function gives $\lambda_{n,j} = \frac{\sqrt{n}j^{j-1}}{j!e^j} \sim \sqrt{\frac{n}{2\pi j^3}}$, for $j, n \rightarrow \infty$, thus the critical phase occurs at $j = \Theta(n^{\frac{1}{3}}): \mathbb{E}(R_{n,j}) \rightarrow \infty$, for $j \ll n^{\frac{1}{3}}$, whereas $\mathbb{E}(R_{n,j}) \rightarrow 0$, for $j \gg n^{\frac{1}{3}}$. In all succeeding examples in this section the critical phase behaviour also occurs at $j = \Theta(n^{\frac{1}{3}})$.

Proof of Theorem 5.1. We consider the description of the problem via the node removal procedure. This immediately yields the following stochastic recurrence for the r.v. $R_{n,j} \stackrel{\mathcal{L}}{=} C_{n,j}^{[v]}$:

$$R_{n,j} \stackrel{\mathcal{L}}{=} R_{K_n,j} + \mathbf{1}_{\{n-K_n=j\}}, \quad \text{for } 1 \leq j \leq n, \quad (5.2)$$

with $R_{n,j} = 0$, for $0 \leq n < j$, and where the r.v. K_n measures the size of the subtree remaining after selecting a random node and removing the subtree rooted at it from a randomly selected size- n Cayley tree.

In the following we will compute the splitting probabilities $p_{n,k} := \mathbb{P}\{K_n = k\}$, with $0 \leq k \leq n-1$, and by doing this we also show that the recurrence (5.2) is indeed valid, i.e., that the subtree T' (let us assume of size $k \geq 1$) remaining after removing the subtree containing the selected node x of a random size- n Cayley tree T is (after an order-preserving relabelling of the nodes) again a random Cayley tree of size k (the so-called *random preservation property* holds).

This can be done by simple combinatorial reasoning. Consider a pair (T, x) of a size- n Cayley tree T and a node $x \in T$. When detaching the subtree rooted at x from T , we obtain a pair (T', T'') of subtrees with T'' containing x and T' the possibly empty remaining subtree. Of course, T' is the empty subtree exactly if x is the root-node of T and thus there are exactly T_n pairs (T, x) yielding $|T'| = k = 0$. Let us now assume that $1 \leq |T'| = k \leq n-1$. After an order-preserving relabelling of T' and T'' with labels $\{1, \dots, k\}$ and $\{1, \dots, n-k\}$, respectively, both subtrees are Cayley trees of size k and $n-k$, respectively. Consider now a particular pair $(\tilde{T}', \tilde{T}'')$ of Cayley trees of size $|\tilde{T}'| = k$ and $|\tilde{T}''| = n-k$, respectively, and let us count the number of pairs (T, x) , with T a size- n Cayley tree and $x \in T$, yielding the pair $(\tilde{T}', \tilde{T}'')$ of subtrees after cutting. By constructing such pairs (T, x) , one obtains that there are exactly $k \binom{n}{k}$ possibilities (k possible ways of attaching \tilde{T}'' to a node in \tilde{T}' and $\binom{n}{k}$ possibilities of distributing the labels $\{1, \dots, n\}$ order-preserving to the subtrees), independent of the chosen pair of trees; thus the random preservation property holds.

Moreover, one obtains that there are exactly $k \binom{n}{k} T_k T_{n-k}$ pairs (T, x) splitting after a cut into a pair (T', T'') of subtrees with respective sizes k and $n-k$, for

$1 \leq k \leq n-1$. Of course, in total there are nT_n pairs (T, x) and thus we get the following result for the splitting probabilities $p_{n,k}$:

$$p_{n,k} = \mathbb{P}\{K_n = k\} = \begin{cases} \frac{k\tilde{T}_k\tilde{T}_{n-k}}{n\tilde{T}_n}, & 1 \leq k \leq n-1, \\ \frac{1}{n}, & k = 0, \end{cases}$$

where we use throughout this section the abbreviation

$$\tilde{T}_n := \frac{T_n}{n!} = \frac{n^{n-1}}{n!}. \quad (5.3)$$

In order to treat the stochastic recurrence (5.2) and to compute the asymptotic behaviour of the (factorial) moments of $R_{n,j}$ we find it appropriate to introduce the generating function

$$F(z, v) := F_j(z, v) = \sum_{n \geq 1} \tilde{T}_n \mathbb{E}(v^{R_{n,j}}) z^n = \sum_{n \geq 1} \sum_{m \geq 0} \tilde{T}_n \mathbb{P}\{R_{n,j} = m\} z^n v^m.$$

Starting with (5.2), straightforward computations (which are omitted here) yield then the following differential equation

$$F_z(z, v) = T(z)F_z(z, v) + \frac{T(z)}{z} + \tilde{T}_j z^{j-1}(v-1) + \tilde{T}_j z^j F_z(z, v)(v-1),$$

where the so-called tree function (the exponential generating function of the number of size- n Cayley trees) appears:

$$T(z) = \sum_{n \geq 1} \tilde{T}_n z^n = \sum_{n \geq 1} \frac{n^{n-1}}{n!} z^n. \quad (5.4)$$

Simple manipulations and using the well-known functional equation of the tree function (which is thus closely related to the Lambert-W function),

$$T(z) = ze^{T(z)},$$

give then the following explicit formula for the derivative of $F(z, v)$ w.r.t. z :

$$F_z(z, v) = \frac{e^{T(z)} + (v-1)\tilde{T}_j z^{j-1}}{1 - T(z) - (v-1)\tilde{T}_j z^j}. \quad (5.5)$$

To get the (factorial) moments of $R_{n,j}$ we use the substitution $w := v-1$ and introduce $\tilde{F}(z, w) := F(z, w+1)$. Extracting coefficients $[w^s]$, $s \geq 1$, from

$$\tilde{F}_z(z, w) = \frac{e^{T(z)} + w\tilde{T}_j z^{j-1}}{1 - T(z) - w\tilde{T}_j z^j} = \frac{1}{z} \left(\frac{1}{1 - T(z) - w\tilde{T}_j z^j} - 1 \right),$$

easily gives

$$[w^s] \tilde{F}_z(z, w) = \frac{\tilde{T}_j^s z^{js-1}}{(1 - T(z))^{s+1}}.$$

We further obtain

$$[z^n w^s] \tilde{F}(z, w) = \frac{1}{n} [z^{n-1} w^s] \tilde{F}_z(z, w) = \frac{\tilde{T}_j^s}{n} [z^{n-j} w^s] \frac{1}{(1 - T(z))^{s+1}}. \quad (5.6)$$

It is not difficult to extract coefficients from (5.6) and stating an explicit formula for $[z^n w^s] \tilde{F}(z, w)$ and the factorial moments of $R_{n,j}$; however, for asymptotic considerations it is easier to use well-known analytic properties of the tree function $T(z)$ and deduce from it the desired asymptotic growth behaviour of $\mathbb{E}(R_{n,j}^s)$. Namely, we use standard applications of so-called singularity analysis, see [24], to transfer the local behaviour of a generating function in a complex neighbourhood of the dominant singularity (i.e., the singularity of smallest modulus; we are here only concerned with functions with a unique dominant singularity) to the asymptotic behaviour of its coefficients. It holds (see, e.g., [24]) that the tree function $T(z)$ has a unique dominant singularity (a branch point) at $z = e^{-1}$, where the function evaluates to $T(e^{-1}) = 1$ and where it admits the following local expansion:

$$T(z) = 1 - \sqrt{2}\sqrt{1 - ez} + \mathcal{O}(1 - ez). \quad (5.7)$$

Thus the function $(1 - T(z))^{-s-1}$ also has a unique dominant singularity at $z = e^{-1}$ with the following local bound:

$$\frac{1}{(1 - T(z))^{s+1}} = \mathcal{O}\left(\frac{1}{(1 - ez)^{\frac{s+1}{2}}}\right).$$

Singularity analysis yields then

$$[z^n] \frac{1}{(1 - T(z))^{s+1}} = \mathcal{O}\left(e^n n^{\frac{s-1}{2}}\right).$$

Therefore, (5.6) yields

$$[z^n w^s] \tilde{F}(z, w) = \mathcal{O}\left(\left(\frac{\tilde{T}_j}{e^j}\right)^s e^n n^{\frac{s-3}{2}}\right).$$

This, together with Stirling's formula for the Gamma function (3.1), shows the following bound for the s -th moments of $R_{n,j}$, which holds uniformly for all $1 \leq j \leq n$:

$$\mathbb{E}(R_{n,j}^s) = \frac{s!}{T_n} [z^n w^s] \tilde{F}(z, w) = \mathcal{O}\left(\left(\frac{\tilde{T}_j}{e^j}\right)^s \frac{e^n n^{\frac{s-3}{2}}}{\tilde{T}_n}\right) = \mathcal{O}\left(\left(\frac{\sqrt{n} \tilde{T}_j}{e^j}\right)^s\right). \quad (5.8)$$

To get the mixed Poisson behaviour for $j = o(n)$ we use the refined expansion

$$\frac{1}{(1 - T(z))^{s+1}} = \frac{1}{2^{\frac{s+1}{2}} (1 - ez)^{\frac{s+1}{2}}} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{1 - ez}}\right)\right),$$

locally around $z = e^{-1}$, which can be obtained from (5.7). Singularity analysis gives then the expansion

$$[z^n] \frac{1}{(1-T(z))^{s+1}} = \frac{e^n n^{\frac{s-1}{2}}}{2^{\frac{s+1}{2}} \left(\frac{s-1}{2}\right)!} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right). \quad (5.9)$$

Thus we get for $j = o(n)$ the stated behaviour of the the s -th factorial moments of $R_{n,j}$:

$$\begin{aligned} \mathbb{E}(R_{n,j}^s) &= \frac{s!}{\tilde{T}_n} [z^n w^s] \tilde{F}(z, w) = \frac{s! \tilde{T}_j^s e^{n-j s} (n-j s)^{\frac{s-1}{2}}}{\tilde{T}_n n 2^{\frac{s+1}{2}} \left(\frac{s-1}{2}\right)!} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n-j s}}\right)\right) \\ &= \left(\frac{\sqrt{n} \tilde{T}_j}{e^j}\right)^s \frac{s! \sqrt{\pi}}{2^{\frac{s}{2}} \left(\frac{s-1}{2}\right)!} \cdot (1 + o(1)) = \left(\frac{\sqrt{n} \tilde{T}_j}{e^j}\right)^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right) \cdot (1 + o(1)), \end{aligned} \quad (5.10)$$

where we used in the final step the duplication formula of the factorials:

$$\left(\frac{s-1}{2}\right)! = \frac{\sqrt{\pi} (s-1)!}{2^{s-1} \left(\frac{s}{2} - 1\right)!}. \quad (5.11)$$

The mixed Poisson limit law with Rayleigh mixing distribution as stated in Theorem 5.1 follows then from (5.8) and (5.10) by applying Lemma 2. \square

5.3. Edge-cutting in Cayley trees

The following prominent edge-cutting procedure for trees is closely related to the node removal procedure considered in Section 5.2. Starting with a tree T one chooses an edge $e \in T$ and removes it from T . After that T decomposes into two subtrees T' and T'' , where we assume that T' contains the original root of T . We discard the subtree T'' and continue the edge-cutting procedure with T' until we have isolated the root-node of the original tree T . This cutting-down procedure has been introduced in [57], where the number of random cuts C_n to isolate the root-node of a random Cayley tree of size n , where in each cutting-step an edge from the remaining tree is chosen uniformly at random, has been studied yielding asymptotic formulæ for the first two moments of C_n . The Rayleigh limiting distribution of C_n for Cayley trees and other families of simply generated trees has been obtained in [61, 62] and in a more general setting by Janson [36]; in particular, for Cayley trees one obtains $\frac{C_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \text{Rayleigh}(1)$. Moreover, in [36] it was shown in general that for Galton-Watson tree families (thus containing Cayley trees) the random variables C_n and $C_n^{[v]}$ (as introduced in Section 5.2) for the edge-cutting procedure and the node-removal procedure, respectively, have the same limiting distribution behaviour. A number of works have analysed the edge-cutting procedure and related processes using the connection between Cayley trees and the so-called continuum random tree, in particular see the

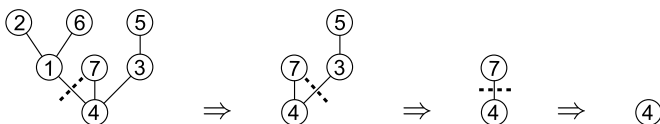


FIG 7. Cutting-down a tree of size 7 with three edge-cuts. The respective sizes of the subtrees cut-off during this procedure are 3, 2, and 1.

work of Addagio-Berry, Broutin and Holmgren [1], and the recent works of Bertoin [5, 9].

In this section we consider a refinement of the r.v. C_n for Cayley trees, namely we study the behaviour of the r.v. $C_{n,j}$, counting the number of subtrees of size j cut-off during the random edge-cutting procedure when starting with a random size- n Cayley tree until the root-node is isolated. Of course, it holds

$$C_n = \sum_{j=1}^{n-1} C_{n,j}.$$

Before continuing we want to remark that an alternative description of the problem can be given via edge-records in edge-labelled trees: given a size- n tree T we first distribute the labels $\{1, \dots, |T| - 1\}$ randomly to the edges of T . An edge-record in T is then an edge $e = (x, y)$, where y is a child of x , with smallest label amongst all edges on the path from the root-node of T to y . Analogous to Section 5.2 (and stated already in [36]) one gets that the r.v. $R_n^{[e]}$ counting the number of edge-records in a random size- n Cayley tree is distributed as C_n , i.e., $R_n^{[e]} \stackrel{\mathcal{L}}{=} C_n$. Moreover, the edge-records e_1, \dots, e_k of an edge-labelled tree T naturally decompose T into the root-node and k edge-record subtrees S_1, \dots, S_k , obtained from T by removing the root-node of T and all edges e_1, \dots, e_k . Again we can introduce the r.v. $R_{n,j}^{[e]}$, which counts the number of edge-record subtrees of size j in a random edge-labelled Cayley tree of size n . It is then immediate to see that $R_{n,j}^{[e]} \stackrel{\mathcal{L}}{=} C_{n,j}$.

In Figure 7 we illustrate the edge-cutting procedure for a particular tree.

In the following theorem we state that $C_{n,j}$ (and thus also $R_{n,j}^{[e]}$) has factorial moments of mixed Poisson type with a Rayleigh mixing distribution. The method of proof is analogous to the one presented in Section 5.2, but due to the less explicit nature of the formulæ occurring, the proof steps are more technical and a bit lengthy.

Theorem 5.2. *The random variable $C_{n,j}$, counting the number of subtrees of size j , which are cut-off during the edge-cutting procedure starting with a random Cayley tree of size n , has, for $n \rightarrow \infty$ and arbitrary $1 \leq j = j(n) \leq n - 1$, asymptotically factorial moments of mixed Poisson type with a Rayleigh mixing distribution X and scale parameter $\lambda_{n,j} = \frac{\sqrt{n}j^{j-1}}{j!e^j}$:*

$$\mathbb{E}(C_{n,j}^s) = \lambda_{n,j}^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right) \cdot (1 + o(1)), \quad \text{for } j = o(n), \quad \mathbb{E}(C_{n,j}^s) = \mathcal{O}(\lambda_{n,j}^s).$$

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{C_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in [0, \infty)$, the random variable $C_{n,j}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$.

Remark 16. According to Theorems 5.1 and 5.2 the r.v. $C_{n,j}$ and $C_{n,j}^{[v]}$ (and thus also $R_{n,j}^{[e]}$ and $R_{n,j}$, for the edge- and vertex-versions of the cutting procedures as considered in Section 5.2-5.3, have the same limiting distribution behaviour. Janson [36] was able to bound the difference between the random variables R_n and $R_n^{[e]}$ (i.e., between the number of node- and edge-records) in a suitable metric and thus to show directly the same limiting behaviour of these r.v. It would be interesting to extend his proof technique to the refined quantities studied here.

Proof of Theorem 5.2. Decomposing a tree according to the first cut of the edge-cutting procedure immediately yields the following stochastic recurrence for the r.v. $C_{n,j}$:

$$C_{n,j} \stackrel{\mathcal{L}}{=} C_{K_n^{[e]},j} + \mathbf{1}_{\{n-K_n^{[e]}=j\}}, \quad \text{for } 1 \leq j < n, \quad (5.12)$$

with $C_{n,j} = 0$, for $1 \leq n \leq j$, and where the r.v. $K_n^{[e]}$ measures the size of the subtree containing the root of the original tree after cutting a random edge from a randomly selected size- n Cayley tree. It is well-known [57] (and can be shown completely analogous to the computations in the proof of Theorem 5.1) that the *random preservation property* of Cayley trees also holds for the edge-cutting procedure (thus implying correctness of (5.12)) and that the splitting probabilities $p_{n,k}^{[e]}$ (i.e., the distribution of $K_n^{[e]}$) are given as follows (we recall the definition $\tilde{T}_n = \frac{n^{n-1}}{n!}$ given in (5.3)):

$$p_{n,k}^{[e]} := \mathbb{P}\{K_n^{[e]} = k\} = \frac{k\tilde{T}_k\tilde{T}_{n-k}}{(n-1)\tilde{T}_n}, \quad \text{for } 1 \leq k \leq n-1 \text{ and } n \geq 2.$$

Again, in order to treat the stochastic recurrence (5.12) we introduce suitable generating functions:

$$G_j(z, v) := \sum_{n \geq 1} \tilde{T}_n \mathbb{E}(v^{C_{n,j}}) z^n = \sum_{n \geq 1} \sum_{m \geq 0} \tilde{T}_n \mathbb{P}\{C_{n,j} = m\} z^n v^m.$$

Straightforward computations give then the differential equation

$$z \frac{\partial}{\partial z} G_j(z, v) - G_j(z, v) = zT(z) \frac{\partial}{\partial z} G_j(z, v) + (v-1)\tilde{T}_j z^{j+1} \frac{\partial}{\partial z} G_j(z, v),$$

where again the tree function $T(z) = \sum_{n \geq 1} \tilde{T}_n z^n$ as defined in (5.4) appears. Solving this linear differential equation yields the following solution satisfying the initial condition $G_j(0, v) = 0$:

$$G_j(z, v) = T(z) \cdot \exp \left(\int_0^z \frac{(v-1)\tilde{T}_j t^{j-1}}{(1-T(t))(1-T(t) - (v-1)\tilde{T}_j t^j)} dt \right). \quad (5.13)$$

To study the moments of $C_{n,j}$ we apply in (5.13) the substitution $w := v - 1$ and introduce the function $\tilde{G}_j(z, w) := G_j(z, w + 1)$, which yields the following expansion w.r.t. w :

$$\tilde{G}_j(z, w) = T(z) \cdot \exp\left(\sum_{\ell \geq 1} w^\ell \alpha_{j,\ell}(z)\right) = T(z) \cdot \left(1 + \sum_{r \geq 1} \frac{1}{r!} \left(\sum_{\ell \geq 1} w^\ell \alpha_{j,\ell}(z)\right)^r\right),$$

where we use the abbreviation

$$\alpha_{j,\ell}(z) := \tilde{T}_j^\ell \cdot \int_0^z \frac{t^{j\ell-1}}{(1-T(t))^{\ell+1}} dt.$$

This leads to the following explicit formula for the generating function of the s -th integer moments of $C_{n,j}$, which will be the starting point of the asymptotic considerations:

$$\begin{aligned} G_j^{[s]}(z) &:= [w^s] \tilde{G}_j(z, w) = \sum_{n \geq 1} \frac{\tilde{T}_n \mathbb{E}(C_{n,j}^s)}{s!} z^n \\ &= T(z) \cdot \sum_{r=1}^s \frac{1}{r!} \sum_{\substack{\ell_1 + \dots + \ell_r = s, \\ \ell_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r \alpha_{j,\ell_q}(z), \quad \text{for } s \geq 1. \end{aligned} \quad (5.14)$$

The following bounds on the growth of the coefficients of the functions appearing, which all can be obtained in a straightforward way by applying standard techniques as singularity analysis or approximating sums by integrals (we omit here some of the details), will play a key rôle in the asymptotic evaluation of the moments. First, it holds, for ℓ arbitrary but fixed and uniformly for $1 \leq j \leq n$ and $n \rightarrow \infty$:

$$[z^n] \alpha_{j,\ell}(z) = \frac{\tilde{T}_j^\ell}{n} [z^{n-j\ell}] \frac{1}{(1-T(z))^{\ell+1}} = \mathcal{O}\left(\frac{\tilde{T}_j^\ell e^n n^{\frac{\ell-3}{2}}}{e^{j\ell}}\right).$$

This implies

$$[z^n] \prod_{q=1}^r \alpha_{j,\ell_q}(z) \Big|_{\substack{\ell_1 + \dots + \ell_r = s, \\ \ell_q \geq 1, 1 \leq q \leq r}} = \mathcal{O}\left(\frac{\tilde{T}_j^s e^n}{e^{js}} \cdot \sum_{\substack{k_1 + \dots + k_r = n, \\ k_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r k_q^{\frac{\ell_q-3}{2}}\right). \quad (5.15)$$

The sum occurring in the latter bound (5.15) can itself be bounded as follows:

$$\sum_{\substack{k_1 + \dots + k_r = n, \\ k_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r k_q^{\frac{\ell_q-3}{2}} = \begin{cases} \mathcal{O}\left(n^{\frac{s}{2} - \frac{r}{2} - 1} (\log n)^Q\right), & \text{for } s > r, \\ \mathcal{O}\left(\frac{(\log n)^{s-1}}{n}\right), & \text{for } s = r, \end{cases}$$

with $Q = |\{q : \ell_q = 1\}|$, thus yielding

$$[z^n] \prod_{q=1}^r \alpha_{j,\ell_q}(z) \Big|_{\substack{\ell_1 + \dots + \ell_r = s, \\ \ell_q \geq 1, 1 \leq q \leq r}} = \begin{cases} \mathcal{O}\left(\frac{\tilde{T}_j^s e^n n^{\frac{s}{2} - \frac{r}{2} - 1} (\log n)^Q}{e^{js}}\right), & \text{for } s > r, \\ \mathcal{O}\left(\frac{\tilde{T}_j^s e^n (\log n)^{s-1}}{e^{js} n}\right), & \text{for } s = r. \end{cases}$$

Therefore, when considering the coefficients of the expression

$$\sum_{r=1}^s \frac{1}{r!} \sum_{\substack{\ell_1+\dots+\ell_r=s, \\ \ell_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r \alpha_{j,\ell_q}(z),$$

the summand $\alpha_{j,s}(z)$ gives the main contribution and implies the following bound, which holds uniformly for $1 \leq j \leq n$ (with s arbitrary but fixed and $n \rightarrow \infty$):

$$[z^n] \sum_{r=1}^s \frac{1}{r!} \sum_{\substack{\ell_1+\dots+\ell_r=s, \\ \ell_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r \alpha_{j,\ell_q}(z) = \mathcal{O}\left(\frac{\tilde{T}_j^s e^n n^{\frac{s-3}{2}}}{e^{js}}\right). \quad (5.16)$$

Now we are in a position to derive the stated bound on the s -th integer moments of $C_{n,j}$. First, by using (5.14) and (5.16) we get for $s \geq 1$:

$$\begin{aligned} \mathbb{E}(C_{n,j}^s) &= \frac{s!}{\tilde{T}_n} [z^n] G_j^{[s]}(z) = \frac{s!}{\tilde{T}_n} [z^n] T(z) \cdot \sum_{r=1}^s \frac{1}{r!} \sum_{\substack{\ell_1+\dots+\ell_r=s, \\ \ell_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r \alpha_{j,\ell_q}(z) \\ &= \mathcal{O}\left(\frac{s!}{\tilde{T}_n} \sum_{k=1}^{n-1} \tilde{T}_k \frac{\tilde{T}_j^s e^{n-k} (n-k)^{\frac{s-3}{2}}}{e^{js}}\right) = \mathcal{O}\left(\frac{\tilde{T}_j^s n^{\frac{3}{2}}}{e^{js}} \sum_{k=1}^{n-1} \frac{(n-k)^{\frac{s-3}{2}}}{k^{\frac{3}{2}}}\right). \end{aligned}$$

Splitting the remaining sum easily gives

$$\begin{aligned} \sum_{k=1}^{n-1} \frac{(n-k)^{\frac{s-3}{2}}}{k^{\frac{3}{2}}} &= \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \frac{(n-k)^{\frac{s-3}{2}}}{k^{\frac{3}{2}}} + \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{n-1} \frac{(n-k)^{\frac{s-3}{2}}}{k^{\frac{3}{2}}} \\ &= \mathcal{O}(n^{\frac{s-3}{2}}) + \begin{cases} \mathcal{O}(n^{\frac{s}{2}-2}), & s \geq 2, \\ \mathcal{O}\left(\frac{\log n}{n^{\frac{3}{2}}}\right), & s = 1 \end{cases} = \mathcal{O}(n^{\frac{s-3}{2}}), \end{aligned} \quad (5.17)$$

thus showing the bound (which holds uniformly for $1 \leq j \leq n$)

$$\mathbb{E}(C_{n,j}^s) = \mathcal{O}\left(\left(\frac{\sqrt{n} \tilde{T}_j}{e^j}\right)^s\right). \quad (5.18)$$

To give the refined asymptotic expansion of $\mathbb{E}(C_{n,j}^s)$, yielding factorial moments of mixed Poisson type, one has to spot and evaluate the main contribution of the coefficients of (5.14) in more detail and to bound the remaining contributions. In order to do this we split

$$\sum_{r=1}^s \frac{1}{r!} \sum_{\substack{\ell_1+\dots+\ell_r=s, \\ \ell_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r \alpha_{j,\ell_q}(z) = \alpha_{j,s}(z) + \underbrace{\sum_{r=2}^s \frac{1}{r!} \sum_{\substack{\ell_1+\dots+\ell_r=s, \\ \ell_q \geq 1, 1 \leq q \leq r}} \prod_{q=1}^r \alpha_{j,\ell_q}(z)}_{=: Q_{j,s}(z)},$$

such that

$$\mathbb{E}(C_{n,j}^s) = \frac{s!}{\tilde{T}_n} [z^n] T(z) \cdot (\alpha_{j,s}(z) + Q_{j,s}(z)). \quad (5.19)$$

Completely analogous to the previous computations one shows for $s \geq 2$ (of course, $Q_{j,1}(z) = 0$):

$$[z^n] Q_{j,s}(z) = \mathcal{O}\left(\frac{\tilde{T}_j^s e^n n^{\frac{s}{2}-2} \log n}{e^{js}}\right),$$

and furthermore (uniformly for $1 \leq j \leq n$) the following bound for the contribution of the remainder:

$$\frac{s!}{\tilde{T}_n} [z^n] T(z) \cdot Q_{j,s}(z) = \mathcal{O}\left(\left(\frac{\sqrt{n} \tilde{T}_j}{e^j}\right)^s \cdot \frac{\log n}{\sqrt{n}}\right). \quad (5.20)$$

Now we consider the term in (5.18) yielding the main contribution, where we assume from now on that $j = o(n)$. We get

$$[z^n] T(z) \alpha_{j,s}(z) = \sum_{k=1}^{n-sj} \tilde{T}_k \tilde{T}_j^s \frac{1}{n-k} [z^{n-k-sj}] \frac{1}{(1-T(z))^{s+1}}. \quad (5.21)$$

We split the summation interval of (5.21) at $k = \lfloor \frac{n}{2} \rfloor$ and consider the contributions separately. Additionally, we only require the already computed asymptotic bounds (5.9) and (5.17). The first part yields

$$\begin{aligned} & \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{T}_k \tilde{T}_j^s \frac{1}{n-k} [z^{n-k-sj}] \frac{1}{(1-T(z))^{s+1}} \\ &= \tilde{T}_j^s \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{T}_k \frac{1}{n-k} \frac{e^{n-k-sj} (n-k-sj)^{\frac{s-1}{2}}}{2^{\frac{s+1}{2}} (\frac{s-1}{2})!} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n-k-sj}}\right)\right) \\ &= \frac{\tilde{T}_j^s e^n n^{\frac{s-3}{2}}}{e^{js} 2^{\frac{s+1}{2}} (\frac{s-1}{2})!} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \frac{\tilde{T}_k}{e^k} \cdot \left(1 + \mathcal{O}\left(\frac{k}{n}\right) + \mathcal{O}\left(\frac{j}{n}\right) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right). \end{aligned}$$

The main contribution comes from

$$\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \frac{\tilde{T}_k}{e^k} = \sum_{k=1}^{\infty} \frac{\tilde{T}_k}{e^k} - \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{\infty} \frac{\tilde{T}_k}{e^k} = T(e^{-1}) + \mathcal{O}\left(\sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{\infty} \frac{1}{k^{\frac{3}{2}}}\right) = 1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right),$$

whereas

$$\mathcal{O}\left(\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \frac{\tilde{T}_k k}{e^k n}\right) = \mathcal{O}\left(\frac{1}{n} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{\sqrt{k}}\right) = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

Thus

$$\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \tilde{T}_k \tilde{T}_j^s \frac{1}{n-k} [z^{n-k-sj}] \frac{1}{(1-T(z))^{s+1}} = \frac{\tilde{T}_j^s e^n n^{\frac{s-3}{2}}}{e^{js} 2^{\frac{s+1}{2}} (\frac{s-1}{2})!} \cdot \left(1 + \mathcal{O}\left(\frac{j}{n}\right) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right). \quad (5.22)$$

The second part yields

$$\begin{aligned}
& \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{n-sj} \tilde{T}_k \tilde{T}_j^s \frac{1}{n-k} [z^{n-k-sj}] \frac{1}{(1-T(z))^{s+1}} \\
&= \mathcal{O}\left(\tilde{T}_j^s \sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{n-sj} \frac{e^k e^{n-k-sj} (n-k-sj)^{\frac{s-1}{2}}}{k^{\frac{3}{2}} (n-k)}\right) = \mathcal{O}\left(\frac{\tilde{T}_j^s e^n}{e^{js} n^{\frac{3}{2}}} \sum_{k=sj}^{\lfloor \frac{n}{2} \rfloor - 1} \frac{(k-sj)^{\frac{s-1}{2}}}{k}\right) \\
&= \mathcal{O}\left(\frac{\tilde{T}_j^s e^n}{e^{js} n^{\frac{3}{2}}} \sum_{k=sj}^{\lfloor \frac{n}{2} \rfloor - 1} k^{\frac{s-3}{2}}\right) = \begin{cases} \mathcal{O}\left(\frac{\tilde{T}_j^s e^n n^{\frac{s-2}{2}}}{e^{js}}\right), & \text{for } s \geq 2, \\ \mathcal{O}\left(\frac{\tilde{T}_j e^n \log n}{e^j n^{\frac{3}{2}}}\right), & \text{for } s = 1. \end{cases} \quad (5.23)
\end{aligned}$$

Starting with (5.21) and combining the contributions (5.22) and (5.23) finally yields

$$[z^n]T(z)\alpha_{j,s}(z) = \left(\frac{\sqrt{n}\tilde{T}_j}{e^j}\right)^s \frac{e^n}{2^{\frac{s+1}{2}} \left(\frac{s-1}{2}\right)! n^{\frac{3}{2}}} \cdot \left(1 + \mathcal{O}\left(\frac{j}{n}\right) + \mathcal{O}\left(\frac{(\log n)^{\delta_{s,1}}}{\sqrt{n}}\right)\right), \quad (5.24)$$

with $\delta_{s,i}$ the Kronecker-delta function. Together with Stirling's formula for the Gamma function we thus obtain from (5.19), (5.20) and (5.24)

$$\mathbb{E}(C_{n,j}^s) = \frac{s!}{\tilde{T}_n} [z^n]G_j^{[s]}(z) = \frac{\sqrt{2\pi}s!}{2^{\frac{s+1}{2}} \left(\frac{s-1}{2}\right)!} \cdot \left(\frac{\sqrt{n}\tilde{T}_j}{e^j}\right)^s \cdot \left(1 + \mathcal{O}\left(\frac{j}{n}\right) + \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right)\right).$$

Finally, applying the duplication formula of the factorials (5.11), we get the stated result for the asymptotic expansion of the factorial moments:

$$\mathbb{E}(C_{n,j}^s) = \left(\frac{\sqrt{n}\tilde{T}_j}{e^j}\right)^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right) \cdot (1 + o(1)), \quad \text{for } j = o(n). \quad (5.25)$$

The mixed Poisson limit law with Rayleigh mixing distribution as stated in Theorem 5.2 follows then from (5.18) and (5.25) by applying Lemma 2. \square

5.4. Parking functions and growth of the initial cluster

Parking functions are objects introduced by Konheim and Weiss [45], which are of interest in combinatorics (e.g., due to connections to various other combinatorial structures as forests, acyclic functions or hyperplane arrangements), see, e.g., [73], and computer science (e.g., due to close relations to hashing variants), see, e.g., [43]. A vivid description of parking functions is as follows: consider a one-way street with n parking spaces numbered from 1 to n and a sequence of n drivers (we will here exclusively deal with the case that the number of parking spaces is equal to the number of drivers) with preferred parking spaces s_1, s_2, \dots, s_n . The drivers arrive sequentially and driver k tries to park at its preferred parking space s_k . If it is free he parks, otherwise he moves further in the allowed direction (thus examining parking spaces $s_k + 1, s_k + 2, \dots$) until

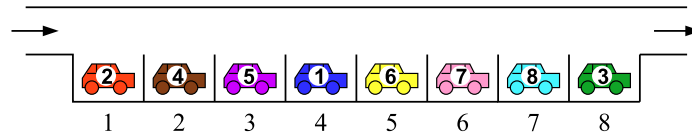


FIG 8. The parking function $(4, 1, 8, 1, 3, 4, 3, 1)$ with the respective parking positions when carrying out the parking procedure.

he finds a free parking space, where he parks; if there is no such parking space he leaves the street without parking. A parking function is then a sequence $(s_1, \dots, s_n) \in \{1, \dots, n\}^n$ such that all drivers are able to park. It has been shown already in [45] that there are exactly $P_n = (n+1)^{n-1}$ parking functions of size n . Figure 8 gives an example of a parking function.

We note that there are many alternative ways of defining parking functions $(s_1, \dots, s_n) \in \{1, \dots, n\}^n$, e.g., via the characterization $|\{j : s_j \leq k\}| \geq k$, for all $1 \leq k \leq n$; however, in what follows the description given above is more intuitive and seems to be advantageous. Namely, we may start with a parking function (s_1, \dots, s_n) and consider the filling of the parking spaces during the *parking procedure*, where at the beginning (step 0) we have an empty street, and where in step k the k -th driver arrives and successfully parks, until (after step n) eventually all parking spaces are occupied. When carrying out the parking procedure, at each time step we may define the initial cluster as the maximum sequence of consecutive occupied parking spaces starting with parking space 1; if parking space 1 is empty we say that the initial cluster is empty. The size of the initial cluster is then simply the number of consecutive occupied parking spaces containing parking space 1. In this section we are interested in the growth of the initial cluster during the parking procedure starting with a random parking function: let the r.v. X_n denote the number of increments of the initial cluster and the refinement $X_{n,j}$ measure the number of increments of amount j of the initial cluster during the parking procedure of a random parking function of size n ; of course $X_n = \sum_{j=1}^{n-1} X_{n,j}$. It turns out that the r.v. X_n and $X_{n,j}$ are closely related to C_n and $C_{n,j}$, respectively, studied in Section 5.3 during the analysis of the edge-cutting procedure of Cayley trees. Figure 9 illustrates the parking procedure and the growth of the initial cluster.

It is well-known [73] that the number P_n of parking functions of size n coincides with the number of rooted labelled forests, i.e., forests of rooted labelled unordered trees, of size n . There are various known bijections between these objects; however, it seems that the following bijective relation between the growth of the initial cluster during the parking procedure and records in forests of rooted labelled trees has not been observed earlier. We mention that a similar bijection has been used by Chassaing and Louchard in [16], where they related the growth of clusters in parking functions with the additive coalescent model for particle coagulation. Analogous to the definition in Section 5.2 a max-record in a forest F of labelled trees is a node $x \in F$, which has the largest label amongst all nodes on the unique path from the root-node of the tree component containing x to x .

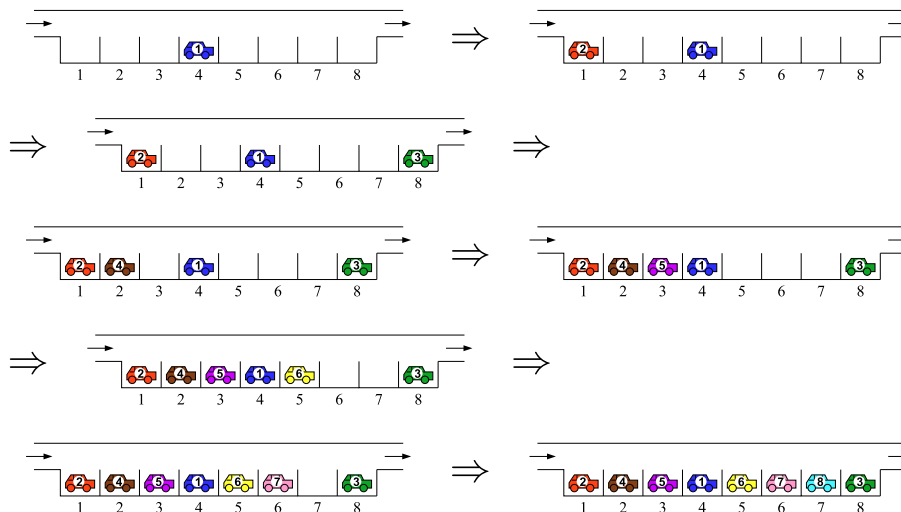


FIG 9. The parking procedure for the parking function $(4, 1, 8, 1, 3, 4, 3, 1)$. There are 6 increments of the initial cluster, 4 increments of amount one, and 2 increments of amount two.

Proposition 4. *There is a bijection, which maps parking functions of size n to forests of rooted labelled unordered trees of size n , such that the number of increments of the initial cluster during the parking procedure of a parking function corresponds to the number of max-records in the forest. Moreover, the number of increments of amount j of the initial cluster during the parking procedure corresponds to the number of max-record subtrees of size j in the forests.*

Proof. Given a parking function $(s_1, \dots, s_n) \in \{1, \dots, n\}^n$ of size n we describe the mapping, i.e., the construction of the corresponding rooted labelled forest F of size n , in an iterative way, which reflects the parking procedure of the n drivers. In order to describe the construction we assume that after step k the first k drivers are parked; then the “parking street” consists of a set of clusters of parking spaces (a cluster is here a maximal sequence of consecutive occupied parking spaces) separated by empty parking spaces. In the construction the k -th driver of the parking function will correspond to the node labelled k in the forest and after step k we will obtain a rooted labelled forest $F^{(k)}$ of size k (with nodes labelled by $\{1, \dots, k\}$). Moreover, the forest $F^{(k)}$ has the property, that each cluster of parking spaces occurring in the parking procedure after step k corresponds to a subset of rooted labelled trees in $F^{(k)}$. It follows the description of the construction of the forest $F := F^{(n)}$:

Step 0 We start with the empty forest $F^{(0)} = \emptyset$.

Step k According to the parking of driver k we distinguish between two cases.

- Driver k parks at his preferred parking space $s := s_k$: let us consider the parking procedure after Step $(k - 1)$ and the cluster of parking

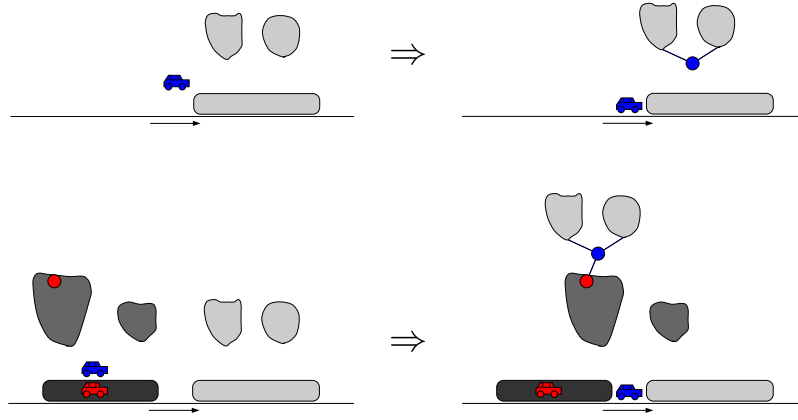


FIG 10. Constructing a forest of labelled rooted trees from a parking function during the parking procedure illustrating the cases, where the preferred parking space of a driver is free (first picture) or occupied (second picture).

spaces starting with parking space $s + 1$ (i.e., the cluster of parking spaces right to the parking space of driver k); if parking space $s + 1$ is empty the cluster is \emptyset . According to the construction this cluster corresponds to a subset G of trees of the forest $F^{(k-1)}$. Let T be the tree rooted at the new node labelled k with G its subtrees. Then the forest $F^{(k)}$ is defined by $F^{(k)} := (F^{(k-1)} \setminus G) \cup T$.

- Driver k cannot park at his preferred parking space s_k , since it is occupied by the ℓ -th driver ($\ell < k$), but parks at the first empty space $s > s_k$: let us consider the parking procedure after Step $(k - 1)$ and the cluster of parking spaces starting with parking space $s + 1$ (i.e., the cluster of parking spaces right to the parking space of driver k); if parking space $s + 1$ is empty the cluster is \emptyset . According to the construction this cluster corresponds to a subset G of trees of the forest $F^{(k-1)}$. Furthermore, let T' be the tree of the forest $F^{(k-1)}$ containing label ℓ (by construction $T' \notin G$). Then, construct the rooted tree T by letting G be the subtrees of the (new) node labelled k and attaching node k to node $\ell \in T'$. Then the forest $F^{(k)}$ is defined by $F^{(k)} := (F^{(k-1)} \setminus (G \cup T')) \cup T$.

Figure 10 illustrates both cases of the bijection.

It follows from the construction that node k is a max-record in the forest F iff the first $k - 1$ drivers occupy all parking spaces left to the parking space s , where the k -th driver has parked, which is equivalent to the event that Step k was an increment of the initial cluster. Moreover, in this case the subtree rooted at node k in $F^{(k)}$ corresponds to a record subtree in the forest F , whose size corresponds then to the amount of increment of the initial cluster. It is not difficult to see that this construction is indeed a bijection from the set of parking functions of

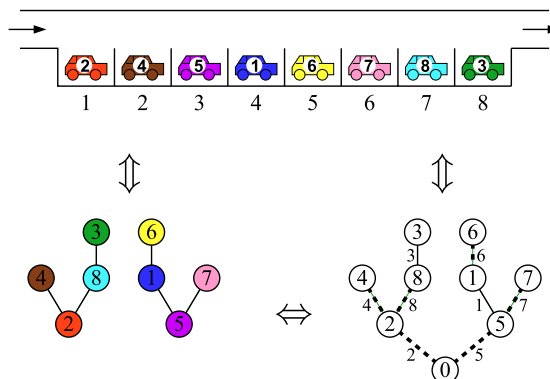


FIG 11. The parking function (4, 1, 8, 1, 3, 4, 3, 1) and the forest of labelled rooted trees obtained by the mapping described in the proof of Proposition 4 as well as the edge-labelled rooted tree described in the proof of Theorem 5.3. The increments of the initial cluster of the parking function correspond to the max-records in the respective forest as well as to the cuts in the respective edge-labelled rooted tree (visualized as dotted lines).

size n to the set of rooted labelled forests of size n , but we omit here to state the inverse mapping (which could be formulated also in an iterative way). \square

In Figure 11 we give a parking function and the corresponding forest of labelled rooted trees under this bijection.

Proposition 4 yields thus a coupling between records in forests of rooted labelled unordered trees and increments in parking functions. This coupling can be extended easily to one between increments in parking functions and edge-cuts to isolate the root-node in Cayley trees.

Theorem 5.3. *The random variable X_n , counting the number of increments of the initial cluster in a random parking function of size n , is equally distributed as the random variable C_n , counting the number of edge-cuts to isolate the root-node in a random Cayley tree of size n , i.e., $X_n \stackrel{\mathcal{L}}{=} C_n$. Moreover, the number of increments of amount j in a random parking function of size n is equally distributed as the number of subtrees of size j cut-off during the edge-cutting procedure when starting with a random size- n Cayley tree, i.e., $X_{n,j} \stackrel{\mathcal{L}}{=} C_{n,j}$, $1 \leq j \leq n - 1$.*

After suitable normalization X_n is asymptotically, for $n \rightarrow \infty$, Rayleigh distributed with parameter 1:

$$\frac{X_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \text{Rayleigh}(1).$$

$X_{n,j}$ has, for $n \rightarrow \infty$ and arbitrary $1 \leq j = j(n) \leq n - 1$, asymptotically factorial moments of mixed Poisson type with a Rayleigh mixing distribution X

and scale parameter $\lambda_{n,j} = \frac{\sqrt{n}j^{j-1}}{j!e^j}$:

$$\mathbb{E}(X_{n,j}^s) = \lambda_{n,j}^s 2^{\frac{s}{2}} \Gamma(\frac{s}{2} + 1) \cdot (1 + o(1)), \text{ for } j = o(n), \quad \mathbb{E}(X_{n,j}^s) = \mathcal{O}(\lambda_{n,j}^s).$$

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{X_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in [0, \infty)$, the random variable $X_{n,j}$ converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \stackrel{\mathcal{L}}{=} MPo(\rho X)$.

Proof. Starting with a labelled forest F of size n we get a rooted labelled tree T of size $n + 1$ by attaching the trees in F as subtrees of the root-node labelled 0. Next we label the edges of T by labels $\{1, \dots, n\}$, where each edge $e = (x, y)$, with x the parent of y , gets the label of the child y . When applying the edge-cutting procedure to the edge-labelled tree T in a way that at each step the edge with largest label is chosen and cut-off, each max-record of the original forest F corresponds to a cut in T and furthermore a record-subtree of size j in F corresponds in T to a cut-off of a branch of size j . Together with Proposition 4 this yields

$$X_n \stackrel{\mathcal{L}}{=} C_n \quad \text{and} \quad X_{n,j} \stackrel{\mathcal{L}}{=} C_{n,j}, \quad 1 \leq j \leq n - 1.$$

The limiting distribution results for X_n and $X_{n,j}$ follow thus from the corresponding results for C_n and $C_{n,j}$ given in [36, 62] and Theorem 5.2, respectively. \square

5.5. Zero contacts in bridges

We consider directed lattice paths from left to right starting at $(0, 0)$ and ending at $(2n, 0)$. At each horizontal unit step we can either go one unit up (step $(1, 1)$) or down (step $(1, -1)$). Such lattice paths are called bridges of length $2n$ starting and ending at height zero, and the steps are stemming from so-called Dyck paths. Of course, such lattice paths are in bijection with lattice paths on a square grid, starting at $(0, 0)$ and ending at (n, n) , with allowed steps $(1, 0)$ (right) and $(0, 1)$ (up). Apparently, there are $B_n = \binom{2n}{n}$ such lattice paths and thus bridges of length $2n$.

Flajolet and Sedgewick [24, Example IX.40, page 707] considered the random variable X_n , counting the number of visits of the x -axis in a random bridge of size $2n$, i.e., the number of k , $1 \leq k \leq n$, with $(2k, 0)$ contained in the bridge, by selecting one of the B_n bridges of length $2n$ uniformly at random. By using a combinatorial decomposition of bridges (the so-called arch decomposition), they have shown that X_n follows asymptotically a Rayleigh distribution, i.e., $\frac{X_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \text{Rayleigh}(\sqrt{2})$.

We consider here a refinement of the r.v. X_n by introducing the r.v. $X_{n,j}$ counting the number of j -visits of the x -axis, where a j -visit is simply a visit

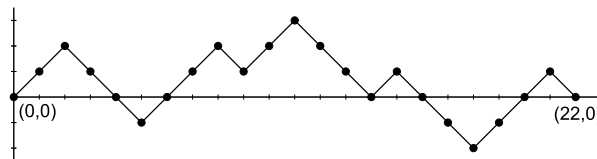


FIG 12. A bridge of length 22 with 6 visits of the x -axis: three 1-visits, two 2-visits, and one 4-visit.

after an excursion of length $2j$, i.e., a return to the x -axis after exactly $2j$, $j \geq 1$, steps. Of course, $X_n = \sum_{j=1}^n X_{n,j}$. Figure 12 illustrates these quantities.

In order to examine the limiting behaviour of $X_{n,j}$ we start with a combinatorial description of the problem using the before mentioned arch decomposition. Let \mathcal{B} be the combinatorial family of bridges of length ≥ 0 and \mathcal{D} be the family of *positive* Dyck path excursions of length ≥ 2 , i.e., Dyck paths of positive length starting and ending on the x -axis, where all points in between are above the x -axis. Analogous, let $\overline{\mathcal{D}}$ be the family of *negative* Dyck path excursions of length ≥ 2 , i.e., Dyck paths of positive length starting and ending on the x -axis, where all points in between are below the x -axis. Then, \mathcal{B} consists of a sequence of positive and negative Dyck path excursions, i.e., it can be described combinatorially by the SEQ construction:

$$\mathcal{B} = \text{SEQ}(\mathcal{D} \dot{\cup} \overline{\mathcal{D}}). \quad (5.26)$$

Furthermore, the family \mathcal{D} can be described formally as

$$\mathcal{D} = \nearrow \text{SEQ}(\mathcal{D}) \searrow. \quad (5.27)$$

Let D_n be the number of positive Dyck path excursions of length $2n$ and $D(z) := \sum_{n \geq 1} D_n z^n$ its generating function; of course, due to symmetry, they coincide with the corresponding quantities for negative Dyck path excursions. Furthermore, let $B(z) := \sum_{n \geq 0} B_n z^n$ be the generating function of the number of bridges B_n of length $2n$. Equations (5.26)-(5.27) immediately yield the following equations for the generating functions,

$$B(z) = \frac{1}{1 - 2D(z)} \quad \text{and} \quad D(z) = \frac{z}{1 - D(z)},$$

with solutions

$$D(z) = \frac{1 - \sqrt{1 - 4z}}{2} \quad \text{and} \quad B(z) = \frac{1}{\sqrt{1 - 4z}}.$$

Of course, by extracting coefficients we reobtain $B_n = \binom{2n}{n}$, whereas the sequence $D_n = \frac{1}{n} \binom{2(n-1)}{n-1}$ is enumerated by the (shifted) Catalan numbers. But more interestingly, the above combinatorial description (5.26)-(5.27) can be extended easily to enumerate a suitably introduced bivariate generating function

of the distribution of $X_{n,j}$:

$$B_j(z, v) := \sum_{n \geq 0} \sum_{m \geq 0} B_n \mathbb{P}\{X_{n,j} = m\} z^n v^m.$$

We get then for $j \geq 1$ the solution

$$B_j(z, v) = \frac{1}{1 - 2(D(z) + (v-1)D_j z^j)} = \frac{1}{\sqrt{1-4z} - 2(v-1)D_j z^j}. \quad (5.28)$$

In order to obtain the factorial moments of $X_{n,j}$ we set $w := v-1$ and introduce the function $\tilde{B}_j(z, w) := B_j(z, w+1)$, thus yielding

$$\tilde{B}_j(z, w) = \frac{1}{\sqrt{1-4z} - 2wD_j z^j}.$$

Extracting coefficients by using standard singularity analysis easily gives

$$\begin{aligned} \mathbb{E}(X_{n,j}^s) &= \frac{s!}{B_n} [z^n w^s] \tilde{B}_j(z, w) = \frac{s!}{B_n} [z^n] \frac{2^s D_j^s z^{js}}{(1-4z)^{\frac{s+1}{2}}} \\ &= \frac{s! 2^s D_j^s}{B_n} [z^{n-js}] \frac{1}{(1-4z)^{\frac{s+1}{2}}} \sim \frac{s! 2^s D_j^s 4^{n-js} (n-js)^{\frac{s-1}{2}}}{B_n \Gamma(\frac{s+1}{2})}. \end{aligned}$$

If $j = o(n)$ we further get

$$\mathbb{E}(X_{n,j}^s) \sim \frac{s! 2^s D_j^s 4^{n-js} n^{\frac{s-1}{2}}}{B_n \Gamma(\frac{s+1}{2})},$$

and, together with $B_n \sim \frac{4^n}{\sqrt{\pi}\sqrt{n}}$ and the duplication formula of the factorials (5.11), this gives

$$\begin{aligned} \mathbb{E}(X_{n,j}^s) &\sim \frac{s! 2^s D_j^s 4^{-js} \sqrt{\pi} n^{\frac{s}{2}}}{\Gamma(\frac{s+1}{2})} = 2^s D_j^s (4^{-j})^s n^{\frac{s}{2}} 2^s \Gamma(\frac{s}{2} + 1) \\ &= \left(\frac{2\sqrt{2}D_j\sqrt{n}}{4^j} \right)^s 2^{\frac{s}{2}} \Gamma(\frac{s}{2} + 1). \end{aligned}$$

Thus, by an application of Lemma 2, we get the following characterization of the limit law of $X_{n,j}$.

Theorem 5.4. *The random variable $X_{n,j}$, counting the number of j -visits of the x -axis in a random bridge of length $2n$, has, for $n \rightarrow \infty$ and arbitrary $1 \leq j = j(n) \leq n$ with $j = o(n)$, asymptotically factorial moments of mixed Poisson type with a Rayleigh mixing distribution X and scale parameter $\lambda_{n,j} = \frac{2\sqrt{2} \binom{2(j-1)}{j-1} \sqrt{n}}{j 4^j}$:*

$$\mathbb{E}(X_{n,j}^s) = (\lambda_{n,j})^s 2^{\frac{s}{2}} \Gamma(\frac{s}{2} + 1) \cdot (1 + o(1)).$$

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{X_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in (0, \infty)$, the random variable $X_{n,j}$ converges in distribution, with convergence of all moments, to $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$.

Moreover, the random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ converges, for $\rho \rightarrow \infty$, after scaling, to its mixing distribution $X: \frac{Y}{\rho} \xrightarrow{\mathcal{L}} X$, with convergence of all moments.

Of course, the result above can be readily adapted to obtain joint distributions for the number of j -visits and the total number of visits of the x -axis as considered by Flajolet and Sedgewick [24]; see also Section 7.1. Our results can be extended to other bridges with different families of steps (see [3]). One can also study this parameter for modified excursions, as considered in [4], and obtain similar results.

5.6. Cyclic points and trees in graphs of random mappings

We call a function $f : [n] \rightarrow [n]$ from the finite set $[n] := \{1, 2, \dots, n\}$ into itself an n -mapping (or an n -mapping function); let us denote by \mathcal{F}_n the set of n -mappings. When selecting one of the n^n n -mappings at random (i.e., if we assume that each of the n^n n -mappings can occur equally likely) one speaks about a random n -mapping. There exists a vast literature (see, e.g., [18, 19, 23] and references therein) devoted to reveal the typical behaviour of important quantities (as, e.g., the number of components, the number of cyclic nodes, etc.) of random n -mappings and the corresponding mapping graphs, respectively.

The *mapping graph*, i.e., the *functional digraph*, of an n -mapping $f \in \mathcal{F}_n$ is the directed graph $G_f = (V, E)$ with set of vertices $V = [n]$ and set of directed edges $E = \{(i, f(i)), i \in [n]\}$. The structure of the mapping graph G_f of an arbitrary mapping function f is well known [18, 23]: the weakly connected components of G_f are cycles of *rooted labelled trees*, i.e., *Cayley trees*, which means that each connected component consists of rooted labelled trees (with edges oriented towards the root nodes) whose root nodes are connected by directed edges such that they are forming a cycle.

This description allows to interpret a mapping f as a set of cycles of labelled trees. Hence, in order to describe the family $\mathcal{F} = \bigcup_{n \geq 0} \mathcal{F}_n$ of all mappings, we can apply the combinatorial constructions SET and CYCLE to the family of rooted labelled trees \mathcal{T} , as introduced and discussed in Section 5.2. This yields the following combinatorial description of the family of mappings \mathcal{F} :

$$\mathcal{F} = \text{SET}(\text{CYCLE}(\mathcal{T})).$$

Hence, the exponential generating function

$$F(z) := \sum_{n \geq 0} n^n \frac{z^n}{n!} = \sum_{n \geq 0} \tilde{F}_n z^n,$$

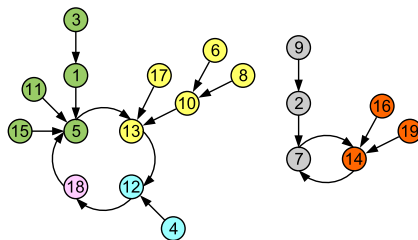


FIG 13. The mapping graph G_f of a 19-mapping function f has 6 cyclic points, i.e., decomposes into 6 trees, one tree of size 1, one tree of size 2, two trees of size 3 and two trees of size 5.

with $\tilde{F}_n := \frac{n^n}{n!}$, of the number of n -mappings satisfies

$$F(z) = \exp\left(\log\left(\frac{1}{1-T(z)}\right)\right) = \frac{1}{1-T(z)}, \quad (5.29)$$

where the tree function $T(z)$ is defined in (5.4). Equation (5.29) suggests the alternative combinatorial description of mappings as sequences of rooted labelled trees: $\mathcal{F} = \text{SEQ}(\mathcal{T})$. This can be justified easily by using an analogue of the canonical cycle representation of permutations: for each cycle of trees order the cycle by starting with the tree having the largest root-label amongst all these trees (let us call it the cycle leader) and then rank the different cycles in descending order of the cycle leaders.

The random variable X_n counting the number of cyclic points in a random n -mapping f (i.e., elements $k \in [n]$, such that there exists a $\ell > 0$ with $k = f^\ell(k)$), which of course coincides with the number of rooted labelled trees in the decomposition of the mapping graph G_f given above, has been analysed by Drmota and Soria [18]. They have shown a Rayleigh limit law: $\frac{X_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \text{Rayleigh}(1)$. We are considering here a refinement of X_n , namely, we introduce the random variables $X_{n,j}$, counting the number of trees of size j occurring in the decomposition of the mapping graph G_f of a random n -mapping $f \in \mathcal{F}_n$; of course, it holds $X_n = \sum_{j=1}^n X_{n,j}$. The quantities considered are visualized in Figure 13.

When introducing a suitable generating function of the distribution of $X_{n,j}$ via $F_j(z, v) := \sum_{n \geq 0} \sum_{m \geq 0} \tilde{F}_n \mathbb{P}\{X_{n,j} = m\} z^n v^m$, the combinatorial decomposition of the family of mappings \mathcal{F} as given above immediately yields an explicit formula for $F_j(z, v)$, $j \geq 1$:

$$F_j(z, v) = \frac{1}{1 - (T(z) - \tilde{T}_j z^j + v \tilde{T}_j z^j)} = \frac{1}{1 - T(z) - (v-1) \tilde{T}_j z^j}.$$

Introducing $w := v - 1$ and $\tilde{F}_j(z, w) := F_j(z, w + 1)$ gives

$$\tilde{F}_j(z, w) = \frac{1}{1 - T(z) - w \tilde{T}_j z^j},$$

from which the factorial moments of $X_{n,j}$ can be obtained easily by extracting coefficients. This can be done completely analogous to the computations in Section 5.2 leading for $j = o(n)$ to the following result:

$$\mathbb{E}(X_{n,j}^s) = \frac{s!}{\tilde{F}_n} [z^n w^s] \tilde{F}_j(z, w) \sim \left(\frac{\tilde{T}_j \sqrt{n}}{e^j} \right)^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right),$$

which, together with Lemma 2, shows the theorem stated below.

Theorem 5.5. *The random variable $X_{n,j}$, counting the number of trees of size j occurring in the decomposition of the mapping graph of a random n -mapping, has, for $n \rightarrow \infty$ and arbitrary $1 \leq j = j(n) \leq n$ with $j = o(n)$, asymptotically factorial moments of mixed Poisson type with a Rayleigh mixing distribution X and scale parameter $\lambda_{n,j} = \frac{j^{j-1} \sqrt{n}}{j! e^j}$:*

$$\mathbb{E}(X_{n,j}^s) = (\lambda_{n,j})^s 2^{\frac{s}{2}} \Gamma\left(\frac{s}{2} + 1\right) \cdot (1 + o(1)).$$

- (i) for $\lambda_{n,j} \rightarrow \infty$, the random variable $\frac{X_{n,j}}{\lambda_{n,j}}$ converges in distribution, with convergence of all moments, to X .
- (ii) for $\lambda_{n,j} \rightarrow \rho \in (0, \infty)$, the random variable $X_{n,j}$ converges in distribution, with convergence of all moments, to $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$.

Moreover, the random variable $Y \stackrel{\mathcal{L}}{=} \text{MPo}(\rho X)$ converges, for $\rho \rightarrow \infty$, after scaling, to its mixing distribution $X: \frac{Y}{\rho} \stackrel{\mathcal{L}}{\rightarrow} X$, with convergence of all moments.

6. Multivariate mixed Poisson distributions

Definition 1 readily extends to multivariate distributions, compare with [20].

Definition 2. *Let (X_1, \dots, X_m) denote a random vector with non-negative components and cumulative distribution function $\mathbf{\Lambda}(\cdot)$ and $\rho_1, \dots, \rho_m > 0$ scale parameters. The discrete random vector (Y_1, \dots, Y_m) with joint probability mass function given by*

$$\mathbb{P}\{Y_1 = \ell_1, \dots, Y_m = \ell_m\} = \frac{\rho_1^{\ell_1} \dots \rho_m^{\ell_m}}{\ell_1! \dots \ell_m!} \int_{(\mathbb{R}^+)^m} X_1^{\ell_1} \dots X_m^{\ell_m} e^{-\sum_{j=1}^m \rho_j X_j} d\mathbf{\Lambda},$$

$\ell_1, \dots, \ell_m \geq 0$, has a multivariate mixed Poisson distribution with mixing distribution (X_1, \dots, X_m) and scale parameters ρ_1, \dots, ρ_m .

Relation (1.6) for the moments of X and Y given there extends to the multivariate case in the following way:

$$\mathbb{E}(Y_1^{s_1} \dots Y_m^{s_m}) = \rho_1^{s_1} \dots \rho_m^{s_m} \mu_{s_1, \dots, s_m}, \quad (6.1)$$

where $\mu_{s_1, \dots, s_m} = \mathbb{E}(X_1^{s_1} \dots X_m^{s_m})$, for $s_1, \dots, s_m \geq 0$; this can readily be seen by a direct computation.

Similar to Proposition 2 we obtain the following result, if the distribution of the random vector (X_1, \dots, X_m) is uniquely determined by the sequence of its (mixed) moments.

Proposition 5. Let $\mathbf{X} = (X_1, \dots, X_m)$ denote a random vector determined by its sequence of mixed moments $(\mu_{s_1, \dots, s_m})_{s_1, \dots, s_m \in \mathbb{N}_0}$, assuming that the moment generating function $\psi(\mathbf{z}) = \mathbb{E}(e^{\mathbf{z}\mathbf{X}})$ of \mathbf{X} exists in a neighbourhood of $\mathbf{0}$ including $-\boldsymbol{\rho}$. Then, the random vector $\mathbf{Y} = (Y_1, \dots, Y_m)$ with mixed factorial moments given by (6.1) has a multivariate mixed Poisson distribution with mixing distribution \mathbf{X} and scale parameters $\rho_1, \dots, \rho_m > 0$. The moment generating function $\varphi(\mathbf{z}) = \mathbb{E}(e^{\mathbf{z}\mathbf{Y}})$ is given by the Stirling transform of $\psi(\mathbf{z})$,

$$\varphi(\mathbf{z}) = (\psi(\mathbf{z})) \Big|_{z_1 = \rho_1(e^{z_1} - 1), \dots, z_m = \rho_m(e^{z_m} - 1)},$$

and the probability mass function of \mathbf{Y} satisfies

$$\mathbb{P}\{Y_1 = \ell_1, \dots, Y_m = \ell_m\} = \sum_{\ell_1 \geq j_1, \dots, \ell_m \geq j_m} \mu_{j_1, \dots, j_m} \prod_{i=1}^m \binom{j_i}{\ell_i} (-1)^{j_i - \ell_i} \frac{\rho_i^{j_i}}{j_i!},$$

for $\ell_1, \dots, \ell_m \geq 0$.

Proof. We proceed similarly to the proof of Proposition 2. Using

$$Y_1^{s_1} \dots Y_m^{s_m} = \sum_{\ell_1=0}^{s_1} \dots \sum_{\ell_m=0}^{s_m} \left(\prod_{i=1}^m Y_i^{s_i} \right)$$

and (2.5), the moment generating function of \mathbf{Y} is readily computed:

$$\begin{aligned} \varphi(\mathbf{z}) &= \sum_{s_1, \dots, s_m \geq 0} \mathbb{E}(Y_1^{s_1} \dots Y_m^{s_m}) \frac{z_1^{s_1} \dots z_m^{s_m}}{s_1! \dots s_m!} \\ &= \sum_{\ell_1, \dots, \ell_m \geq 0} \mu_{\ell_1, \dots, \ell_m} \prod_{i=1}^m \left(\rho_i (e^{z_i} - 1) \right)^{\ell_i}. \end{aligned}$$

Moreover, it coincides with the Stirling transform of $\psi(\mathbf{z})$. This proves that the random vector has a multivariate mixed Poisson law, since $\varphi(\mathbf{z})$ is analytic in a neighbourhood of $\mathbf{0}$. Moreover, the probability mass function is obtained according to

$$\mathbb{P}\{Y_1 = \ell_1, \dots, Y_m = \ell_m\} = \frac{\rho_1^{\ell_1} \dots \rho_m^{\ell_m}}{\ell_1! \dots \ell_m!} \cdot \left(\frac{\partial^{\sum_{k=1}^m \ell_k}}{\partial z_1^{\ell_1} \dots \partial z_m^{\ell_m}} \varphi(\mathbf{z}) \right) \Big|_{\mathbf{z} = -\boldsymbol{\rho}}.$$

□

Moreover, we note that the basic limit theorem for the univariate case given in Lemma 2 can be readily extended to limit laws for random vectors.

7. Outlook and extensions

7.1. Mixed Poisson distributions in Analytic Combinatorics - compositions

Looking at the examples in Sections 5.2 and 5.5 leading to mixed Poisson laws with Rayleigh mixing distribution, and the example in Section 3.1 from [51], it

is desirable to find a unifying combinatorial scheme leading to mixed Poisson distributions. The generating functions appearing in the examples mentioned beforehand indicate how to do so: we can use the critical compositions (see Flajolet and Sedgewick [24, Proposition IX.24, page 712], based on the pioneering works of Soria et al. [2, 18, 19, 25, 26, 27]. Assumed that generating functions $G(z)$ and $H(z)$ are the counting series of certain combinatorial families. We are interested in compositions of generating functions of the form $G(H(z))$. Combinatorially, this amounts to a substitution between structures of the form $\mathcal{F} = \mathcal{G} \circ \mathcal{H}$. We measure the size of the so-called core X_n , and additionally taking into account the contribution of parts of size j to the core:

$$F(z, u, v) = G(u(H(z) - (v-1)H_j z^j)).$$

Here the variable u marks as usual the total size of the so-called core X_n ,

$$\mathbb{P}\{X_n = m\} = \frac{[z^n u^k]F(z, u, 1)}{[z^n]G(H(z))},$$

and the new variable v marks the contribution of parts of size j measured by the random variable $X_{n,j}$ to the core,

$$\mathbb{P}\{X_{n,j} = m\} = \frac{[z^n v^j]F(z, 1, v)}{[z^n]G(H(z))},$$

such that $X_n = \sum_{j=1}^n X_{n,j}$. Using the semi-large power theorem stated in [24, Theorem IX.16, page 709], one can study this j -part core $X_{n,j}$. More generally, it is desirable to study the joint distributions $(X_n; X_{n,j_1}, \dots, X_{n,j_k})$ via

$$F(z, u, \mathbf{v}) = G\left(u(H(z) - \sum_{\ell=1}^k (v_\ell - 1)H_{j_\ell} z^{j_\ell})\right).$$

We will report on our findings on this refined analysis of compositions elsewhere [46].

7.2. Extensions and open problems

Several of the results presented in this work can be extended to a multivariate analysis involving a random vector $\mathbf{X}_{n,\mathbf{j}} = (X_{n,j_1}, \dots, X_{n,j_r})$ by its mixed moments. It should be possible, at least for some of the examples presented, to use mixed Poisson approximation techniques to derive distances, i.e., the total variation distance, between the random variables of interest and the corresponding mixed Poisson distributions, refining the results stated in this work.

A natural question is to ask for a direct explanation of the critical growth rates of the second parameter, say j , when the limit laws of the random variables $X_{n,j}$ change from the law of the mixing distribution X to the mixed Poisson distributions $\text{MPo}(\rho X)$.

Acknowledgements

The authors thank Svante Janson for very helpful remarks on mixed Poisson distributions, in particular regarding notation and the correct approach to Lemma 1. Furthermore, the first author warmly thanks Walter Kuba for several motivating discussions and for providing a concise argument used in the series related part of Lemma 1.

The authors thank the referee for valuable remarks improving the presentation of this article.

References

- [1] L. Addario-Berry, N. Broutin, C. Holmgren, Cutting down trees with a Markov chainsaw, *Annals of Applied Probability* 24, 2297–2339, 2014. [MR3262504](#)
- [2] C. Banderier, P. Flajolet, G. Schaeffer, M. Soria, Random maps, coalescing saddles, singularity analysis and Airy phenomena, *Random Structures & Algorithms* 19, 194–246, 2001. [MR1871555](#)
- [3] C. Banderier and P. Flajolet, Basic analytic combinatorics of directed lattice paths. *Theoretical Computer Science* 281, 37–80, 2002. [MR1909568](#)
- [4] C. Banderier and M. Wallner, *Some reflections on directed lattice paths*, *DMTCS Proceedings Series*, Proceedings of AofA 2014, to appear, 2014.
- [5] J. Bertoin, Fires on trees, *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* 48, no. 4, 909–921, 2012. [MR3052398](#)
- [6] J. Bertoin, Sizes of the largest clusters for supercritical percolation on random recursive trees. *Random Structures & Algorithms* 44, 29–44, 2014. [MR3143589](#)
- [7] J. Bertoin, The cut-tree of large recursive trees, *Annales de l'Institut Henri Poincaré*, to appear. [MR3335011](#)
- [8] E. Baur and J. Bertoin, Cutting edges at random in large recursive trees, Preprint, available at <http://hal.archives-ouvertes.fr/docs/01/00/31/51/PDF/cutting-edges.pdf>. [MR3332709](#)
- [9] J. Bertoin and G. Miermont, The cut-tree of large Galton-Watson trees and the Brownian CRT, *Annals of Applied Probability* 23, 1469–1493, 2013. [MR3098439](#)
- [10] Z.-D. Bai, F. Hu and L.-X. Zhang, Gaussian approximation theorems for urn models and their applications, *Annals of Applied Probability* 12, 1149–1173, 2002. [MR1936587](#)
- [11] F. Bergeron, P. Flajolet and B. Salvy, Varieties of Increasing Trees, *Lecture Notes in Computer Science* 581, 24–48, 1992. [MR1251994](#)
- [12] M. Bernstein and N. J. A. Sloane, Some canonical sequences of integers, *Linear Algebra and its Applications* 226/228, 57–72, 1995. [MR1344554](#)
- [13] F. Brenti, Unimodal, log-concave, and Pólya frequency sequences in combinatorics, *Memoirs Amer. Math. Soc.* 81, no. 413, 1989. [MR0963833](#)

- [14] F. Brenti, Hilbert polynomials in combinatorics, *J. Algebraic Combinatorics* 7, 127–156, 1998. [MR1609885](#)
- [15] T. Carleman, Sur le problème des moments, *Acad. Sci. Paris* 174, 1680–1682, 1922.
- [16] P. Chassaing and G. Louchard, Phase transition for parking blocks, Brownian excursion and coalescence, *Random Structures & Algorithms* 21, 76–119, 2002. [MR1913079](#)
- [17] J. H. Curtiss, A note on the theory of moment generating functions, *Annals of Math. Stat.* 13, 430–433, 1942. [MR0007577](#)
- [18] M. Drmota and M. Soria, Images and preimages in random mappings, *SIAM Journal on Discrete Mathematics* 10, 246–269, 1997. [MR1445035](#)
- [19] M. Drmota and M. Soria. Marking in combinatorial constructions: Generating functions and limiting distributions, *Theoretical Computer Science* 144, 67–99, 1995. [MR1337754](#)
- [20] A. Ferrari, G. Letacy and J.-Y. Tournier, Multivariate mixed Poisson distributions, *EUSIPCO-04*, Vienna, Austria, 2004.
- [21] P. Flajolet, P. Dumas and V. Puyhaubert, Some exactly solvable models of urn process theory, *Discrete Mathematics and Theoretical Computer Science*, vol. AG, 59–118, 2006, in “Proceedings of Fourth Colloquium on Mathematics and Computer Science”, P. Chassaing Editor. [MR2509623](#)
- [22] P. Flajolet, J. Gabarró and H. Pekari, Analytic urns, *Annals of Probability* 33, 1200–1233, 2005. [MR2135318](#)
- [23] P. Flajolet and A. Odlyzko, Random mapping statistics, in: *Advances in cryptology – EUROCRYPT ’89*, Lecture Notes in Computer Science 434, 329–354, Springer, Berlin, 1990. [MR1083961](#)
- [24] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, Cambridge, UK, 2009. [MR2483235](#)
- [25] P. Flajolet and M. Soria, Gaussian limiting distributions for the number of components in combinatorial structures, *J. Combinatorial Theory, Series A* 53, 165–182, 1990. [MR1041444](#)
- [26] P. Flajolet and M. Soria, The cycle construction, *SIAM Journal on Discrete Mathematics* 4, 58–60, 1991. [MR1090289](#)
- [27] P. Flajolet and M. Soria, General combinatorial schemas: Gaussian limit distributions and exponential tails, *Discrete Mathematics* 114, 159–180, 1993. [MR1217750](#)
- [28] P. Flajolet and J. Vitter, Average case analysis of algorithms and data structures, in *Handbook of Theoretical Computer Science*, 431–524, Elsevier, Amsterdam, 1990. [MR1127175](#)
- [29] I. Gessel and R. P. Stanley, Stirling polynomials, *Journal of Combinatorial Theory, Series A* 24, 24–33, 1978. [MR0462961](#)
- [30] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Second Edition. Reading, Massachusetts: Addison-Wesley, 1994. [MR1397498](#)
- [31] R. Grübel and N. Stefanoski, Mixed Poisson approximation of node depth distributions in random binary search trees, *Annals of Applied Probability* 15, 279–297, 2005. [MR2115044](#)
- [32] A. Gut, On the moment problem, *Bernoulli* 8, 407–421, 2002. [MR1913113](#)

- [33] H.-K. Hwang and R. Neininger, Phase change of limit laws in the quicksort recurrence under varying toll functions, *SIAM Journal on Computing* 31, 1687–1722, 2002. [MR1954876](#)
- [34] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns, *Stochastic processes and applications* 110, 177–245, 2004. [MR2040966](#)
- [35] S. Janson, Random records and cuttings in complete binary trees, *Mathematics and Computer Science III, Algorithms, Trees, Combinatorics and Probabilities*, M. Drmota, P. Flajolet, D. Gardy, B. Gittenberger (eds.), 241–253, Birkhäuser, Basel, 2004. [MR2090513](#)
- [36] S. Janson, Random cutting and records in deterministic and random trees, *Random Structures & Algorithms* 29, 139–179, 2006. [MR2245498](#)
- [37] S. Janson, Limit theorems for triangular urn schemes, *Probability Theory and Related Fields* 134, 417–452, 2005. [MR2226887](#)
- [38] S. Janson, Plane recursive trees, Stirling permutations and an urn model, Proceedings, Fifth Colloquium on Mathematics and Computer Science (Blaubeuren, 2008), *DMTCS Proceedings AI*, 541–548, 2008. [MR2508813](#)
- [39] S. Janson, Moments of Gamma type and the Brownian supremum process area, *Probability Surveys*, 7, 1–52, 2010. [MR2645216](#)
- [40] S. Janson, M. Kuba and A. Panholzer, Generalized Stirling permutations, families of increasing trees and urn models, *Journal of Combinatorial Theory, Series A* 118, 94–114, 2011. [MR2737187](#)
- [41] N. L. Johnson, S. Kotz and A. W. Kemp, *Univariate Discrete Distributions*, 2. Edition, New York, John Wiley, 1992. [MR1224449](#)
- [42] D. Karlis and E. Xekalaki, Mixed Poisson Distributions, *International Statistical Review* 73, 35–58, 2005.
- [43] D. E. Knuth, *The Art of Computer Programming, Volume 3: Sorting and Searching*, second edition, Addison-Wesley, Reading, 1998. [MR3077154](#)
- [44] L. M. Koganov, Universal bijection between Gessel-Stanley permutations and diagrams of connections of corresponding ranks, *Uspekhi Mat. Nauk* 51, no. 2(308), 165–166, 1996; English translation in *Russian Math. Surveys* 51, no. 2, 333–335, 1996. [MR1401550](#)
- [45] A. G. Konheim and B. Weiss, An occupancy discipline and applications, *SIAM Journal on Applied Mathematics* 14, 1266–1274, 1966.
- [46] M. Kuba, A note on Critical Compositions and Mixed Poisson distributions, manuscript.
- [47] M. Kuba and A. Panholzer, Descendants in increasing trees, *Electronic Journal of Combinatorics* 13 (1), Paper 8, 2006. [MR2200536](#)
- [48] M. Kuba and A. Panholzer, Analysis of label-based parameters in increasing trees, *Discrete Mathematics and Theoretical Computer Science*, Proceedings AG, 321–330, 2006. [MR2509642](#)
- [49] M. Kuba and A. Panholzer, On the degree distribution of the nodes in increasing trees, *Journal of Combinatorial Theory, Series A* 114, 597–618, 2007. [MR2319165](#)
- [50] M. Kuba and A. Panholzer, Isolating nodes in recursive trees, *Aequationes Mathematicae* 76, 258–280, 2008. [MR2461893](#)

- [51] M. Kuba and A. Panholzer, Analysis of statistics for generalized Stirling permutations, *Combinatorics, Probability and Computing* 20, 875–910, 2011. [MR2847273](#)
- [52] M. Kuba and A. Panholzer, On death processes and urn models, *Discrete Mathematics and Theoretical Computer Science*, in: “Proceedings of the 23rd International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2012)”, Proceedings AQ, 29–42, 2012. [MR2957318](#)
- [53] M. Kuba and A. Panholzer, Limiting distributions for a class of diminishing urn models, *Advances in Applied Probability* 44, 1–31, 2012. [MR2951548](#)
- [54] M. Kuba and A. Panholzer, Isolating nodes in recursive trees, *Online Journal of Analytic Combinatorics* 9, 2014, 26 pages. [MR3259471](#)
- [55] M. Loève, *Probability Theory I*, 4th Edition, Springer-Verlag, New York, 1977. [MR0651017](#)
- [56] H. Mahmoud and R. Smythe, On the distribution of leaves in rooted subtrees of recursive trees, *Annals of Applied Probability* 1, 406–418, 1991. [MR1111525](#)
- [57] A. Meir and J. W. Moon, Cutting down random trees, *Journal of the Australian Mathematical Society* 11, 313–324, 1970. [MR0284370](#)
- [58] A. Meir and J. W. Moon, Cutting down recursive trees, *Mathematical Biosciences* 21, 173–181, 1974.
- [59] J. C. Massé, and R. Theodorescu, Neyman type A distribution revisited, *Statistica Neerlandica*, Volume 59, Issue 2, 206–213, 2005. [MR2153615](#)
- [60] J. Neyman, On a new class of “contagious” distributions applicable in entomology and bacteriology, *Annals of Mathematical Statistics* 10, 35–57, 1939.
- [61] A. Panholzer, Non-crossing trees revisited: cutting down and spanning subtrees, *Discrete Random Walks*, C. Banderier and C. Krattenthaler (eds.), *Discrete Mathematics and Theoretical Computer Science*, Proceedings AC, 265–276, 2003. [MR2042393](#)
- [62] A. Panholzer, Cutting down very simple trees, *Quaestiones Mathematicae* 29, 211–228, 2006. [MR2233368](#)
- [63] A. Panholzer and H. Prodinger, The level of nodes in increasing trees revisited, *Random Structures & Algorithms* 31, 203–226, 2007. [MR2343719](#)
- [64] A. Panholzer and G. Seitz, Limiting distributions for the number of inversions in labelled tree families, *Annals of Combinatorics* 16, 847–870, 2012. [MR3000449](#)
- [65] A. Panholzer and G. Seitz, Ancestors and descendants in evolving k -tree models, *Random Structures & Algorithms* 44, 465–489, 2014. [MR3214201](#)
- [66] S. K. Park, The r -multipermutations, *Journal of Combinatorial Theory, Series A* 67, 44–71, 1994. [MR1280598](#)
- [67] S. K. Park, Inverse descents of r -multipermutations, *Discrete Mathematics* 132, 215–229, 1994. [MR1297383](#)
- [68] S. K. Park, P -partitions and q -Stirling numbers, *Journal of Combinatorial Theory, Series A* 68, 33–52, 1994. [MR1295782](#)

- [69] J. Pitman, Exchangeable and partially exchangeable random partitions, *Probab. Theory Relat. Fields* 102, 145–158, 1995. [MR1337249](#)
- [70] N. Pouyanne, An algebraic approach to Pólya processes, *Annales de l'Institut Henri Poincaré* 44, 293–323, 2008. [MR2446325](#)
- [71] N. Privault, Generalized Bell polynomials and the combinatorics of Poisson central moments, *Electronic Journal of Combinatorics* 18, Paper 54, 2011. [MR2776830](#)
- [72] V. Puyhaubert, *Modèles d'urnes et phénomènes de seuils en combinatoire analytique*, Ph.D. thesis, École Polytechnique, Palaiseau, 2005.
- [73] R. Stanley, *Enumerative Combinatorics, Volume I & II*, Cambridge University Press, 1997 & 1999.
- [74] C. Su, Q. Feng and Z. Hu, Uniform recursive trees: Branching structure and simple random downward walk, *Journal of Mathematical Analysis and Applications* 315, 225–243, 2006. [MR2196543](#)
- [75] Eric W. Weisstein, “Stirling Transform”, From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/StirlingTransform.html>
- [76] H. S. Wilf, *Generatingfunctionology*, Second Edition, Academic Press, 1992. [MR1277813](#)
- [77] G. Willmot, Mixed compound Poisson distributions, *ASTIN Bulletin* 16, 59–80, 1986.