ON MOMENT SEQUENCES AND MIXED POISSON DISTRIBUTIONS

MARKUS KUBA AND ALOIS PANHOLZER

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 simple limit theorem based on expansions of factorial 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 analyze 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-oriented 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.

1. INTRODUCTION

In combinatorics the Stirling transform of a given sequence $(a_s)_{s \in \mathbb{N}}$, see [7, 68], is the sequence $(b_s)_{s \in \mathbb{N}}$, with elements given by

$$b_s = \sum_{k=1}^{s} \binom{s}{k} a_k, \quad s \geq 1.$$  

(1)

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

$$a_s = \sum_{k=1}^{s} (-1)^{s-k} \binom{s}{k} b_k, \quad s \geq 1.$$  

(2)

Here $\binom{s}{k}$ 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 [66] or [25], and $\binom{n}{m}$ denotes the unsigned Stirling numbers of the first kind, counting the number of permutations of $n$ elements with $m$ cycles [25]. These numbers appear as
coefficients in the expansions
\[ x^s = \sum_{k=0}^{s} \binom{s}{k} x^k, \quad x^s = \sum_{k=0}^{s} (-1)^{s-k} \binom{s}{k} x^k, \quad (3) \]
relating ordinary powers \( x^s \) to the so-called falling factorials\( x^s = x(x-1) \ldots (x-(s-1)), \ s \in \mathbb{N}_0 \). On the level of exponential generating functions \( A(z) = \sum_{s \geq 0} a_s z^s/s! \) and \( B(z) = \sum_{s \geq 0} b_s z^s/s! \), the Stirling transform and the relations (1) and (2) turn into
\[ B(z) = A(e^z - 1), \quad A(z) = B(\log(1 + z)). \quad (4) \]
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
\[ b_s = \sum_{k=1}^{s} \rho^k \binom{s}{k} a_k, \quad \text{such that } a_s = \frac{1}{\rho^s} \sum_{k=1}^{s} (-1)^{s-k} \binom{s}{k} b_k, \quad s \geq 1. \quad (5) \]
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 \textit{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 the random variable \( Y \), given its sequence of \textit{factorial moments} \( \mathbb{E}(Y^s) = \mathbb{E}(Y(Y-1) \ldots (Y-(s-1))) \), which are determined by the moments of \( X \),
\[ \mathbb{E}(Y^s) = \rho^s \mathbb{E}(X^s) = \rho^s \mu_s, \quad s \geq 1, \quad (6) \]
where \( \rho > 0 \) denotes an auxiliary scale parameter. Moreover, we discuss relations between the moment generating functions \( \psi(z) = \mathbb{E}(e^{zX}), \ \varphi(z) = \mathbb{E}(e^{zY}) \) of \( X \) and \( Y \), respectively.

1.1. Motivation. Our main motivation to study random variables with a given sequences of factorial moments (6) stems from the analysis of combinatorial structures. In many cases, amongst others the analysis of inversions in labelled tree families [56], stopping times in urn models [47, 45], node degrees in increasing trees [44], block sizes in \( k \)-Stirling permutations [45], descendants in increasing trees [42], ancestors and descendants in evolving k-tree models [57], pairs of random variables \( X \) and \( Y \) arise as limiting distributions for certain parameters of interests 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 (6). An open problem was to understand in more detail the nature of the random variable \( Y \). In [56, 47] 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; we will also add a few 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 [53, 70, 52, 37, 15]. In the analysis of
random discrete structures mixed Poisson distributions have been used mainly in the context of Poisson approximation, see e.g. [26]. 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 $x^s = x(x-1)\ldots(x-(s-1))$ for the falling factorials, and $x^s = x(x+1)\ldots(x+s-1)$ for the rising factorials.\(^1\) Moreover, we denote with $\{\binom{s}{k}\}$ the Stirling numbers of the second kind. We use the notation $U \overset{L}{=} V$ for the equality in distribution of random variables $U$ and $V$, and $U_n \overset{L}{\to} V$ denotes the converge in distribution of a sequence of random variables $U_n$ to $V$. The indicator variable of the event $A$ is denoted by $1_A$. 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. Furthermore, we denote with $E_v$ the evaluation operator of the variable $v$ to 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 distributions with Rayleigh mixing distribution; in particular, we discuss a node isolation in Cayley trees, directed lattice paths and zero contacts, 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 obtain a random variable $Y$ with prescribed sequence of factorial moments, given according to Equation (6) by $E(Y^s) = \mu^s$, a first ansatz would be the following. Let $Y$ denote a discrete random variable supported on the non-negative integers, and $p(v)$ its probability generating function,

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

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

\[ E(Y^s) = \sum_{\ell \geq 0} \ell^s \mathbb{P}\{Y = \ell\} = E_v D_v^s p(v), \quad s \geq 0. \]  \hspace{1cm} (7)

\(^1\)The notation $x^s$ and $x^s$ 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.
Consequently, we can describe the probability mass function of a random variable $Y$ as follows:

$$p(v) = \sum_{s \geq 0} E(Y^s) \left(\frac{(v-1)^s s!}{s!}\right) = \sum_{s \geq 0} (v-1)^s \sum_{j \geq 0} v^j \sum_{s \geq j} \binom{s}{j} (-1)^{s-j} \frac{\mu^s \rho^s}{s!}.$$  

This implies that

$$P\{Y = \ell\} = \left[v^\ell\right] p(v) = \sum_{s \geq \ell} \binom{s}{\ell} (-1)^{s-\ell} \frac{\mu^s \rho^s}{s!}, \quad \ell \geq 0. \tag{8}$$

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 $\Lambda(.)$, then the discrete random variable $Y$ with probability mass function given by

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

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

The boundary case $X \overset{\ell}{=} 0$ 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: $P\{Y = \ell\} = \frac{1}{\ell!} E(X^\ell e^{-X})$.

One often encounters a slightly different definition, which includes a scale parameter $\rho \geq 0$:

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

This corresponds to a scaling of the mixing distribution, $Y \overset{\ell}{=} \text{MPo}(\rho X)$. Here and throughout this work we call $Y \overset{\ell}{=} \text{MPo}(\rho X)$ a mixed Poisson distributed random variable with mixing distribution $X$ and scale parameter $\rho$.

#### Example 1.

The ordinary Poisson distribution $Y \overset{\ell}{=} \text{Po}(\rho)$ with parameter $\rho > 0$,

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

arises as a mixed Poisson distribution with degenerate mixing distribution $X \overset{\ell}{=} 1$.

#### Example 2.

The negative binomial distribution $Y \overset{\ell}{=} \text{NegBin}(r, p)$ with parameters $p \in [0, 1)$ and $r > 0$,

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

arises as a mixed Poisson distribution with mixing distribution $X \overset{\ell}{=} \text{Gamma}(r, \theta)$ scaled by $\rho \geq 0$, such that the parameters $\theta$ and $\rho$ 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 $\text{Geom}(p) = \text{NegBin}(1, p)$.

**Example 3.** A Rayleigh distributed r.v. $X \overset{\text{L}}{=} \text{Rayleigh}(\sigma)$ with parameter $\sigma$ 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:
\[
\mathbb{E}(X^s) = \sigma^s 2^s \Gamma\left(\frac{s}{2} + 1\right).
\]
A discrete random variable $Y$ with probability mass function
\[
\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} \int_0^\infty x^\ell e^{-\rho x - \frac{x^2}{2}} dx, \quad \ell \geq 0,
\]
arises as a mixed Poisson distribution $Y \overset{\text{L}}{=} \text{MPo}(\rho X)$ with mixing distribution $X \overset{\text{L}}{=} \text{Rayleigh}(1)$ and scale parameter $\rho$. We call $Y$ a Poisson-Rayleigh distribution with parameter $\rho$. 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 $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$:
\[
\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} e^{\frac{\ell+1}{2}} \sum_{i=0}^{\ell+1} \binom{\ell+1}{i} (-\rho)^{\ell+1-i} \frac{i+1}{i} \Gamma\left(\frac{i+1}{2}, \frac{\rho^2}{2}\right).
\]

**Example 4.** The Neyman Type A Distribution is a discrete probability distribution often used in biology and ecology [53, 52]. It is a mixed Poisson distribution with mixing distribution $X \overset{\text{L}}{=} \text{Po}(\lambda)$ given by an (ordinary) Poisson distribution with parameter $\lambda$, scaled by $\rho$:
\[
\mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} e^{-\lambda+\rho e^{-\rho}} \sum_{m=0}^\ell \frac{\lambda^m}{m!} = \frac{\rho^\ell}{\ell!} e^{-\lambda+\rho e^{-\rho}} \sum_{j=0}^\ell \binom{\ell}{j} (\rho e^{-\rho})^j.
\]

For a very comprehensive list of examples of mixed Poisson distributions we refer the reader to the article of Willmot [70]. Since by (3) the factorial moments $\mathbb{E}(Y^s)$ 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 $Y \overset{\text{L}}{=} \text{MPo}(\rho X)$ 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, $\mathbb{E}(Y^s) = \rho^s \mathbb{E}(X^s), \ s \geq 1$.

(b) The power moments of $Y$ and $X$ are related by the generalized Stirling transform with parameter $\rho$, and its inverse, respectively:
\[
\mathbb{E}(Y^s) = \sum_{j=0}^s \binom{s}{j} \rho^j \mathbb{E}(X^j), \quad \mathbb{E}(X^s) = \frac{1}{\rho^s} \sum_{j=0}^s (-1)^{s-j} \binom{s}{j} \mathbb{E}(Y^j).
\]
Similarly, the cumulants of \( Y \) and \( X \) are related by the generalized Stirling transform with parameter \( \rho \), 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:

\[
\varphi(z) = \psi\left(\rho(e^z - 1)\right), \quad \psi(z) = \varphi\left(\log\left(\frac{z}{\rho} + 1\right)\right). \tag{9}
\]

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

\[
Y \overset{\xi}{=} \text{MPo}(\rho_1 X_1 \oplus \rho_2 X_2).
\]

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

\[
\mathbb{E}(Y^z) = \sum_{\ell \geq 0} \ell^z \mathbb{P}(Y = \ell) = \sum_{\ell \geq s} \ell^z \frac{\rho^\ell}{\ell!} \int_{\mathbb{R}^+} X^s e^{-\rho X} d\Lambda
\]

\[
= \rho^s \sum_{\ell \geq s} \frac{\rho^{\ell-s}}{(\ell-s)!} \mathbb{E}(X^\ell e^{-\rho X}) = \rho^s \mathbb{E}(X^s e^{-\rho X} \sum_{\ell \geq 0} \frac{\rho^{\ell} X^\ell}{\ell!}) = \rho^s \mathbb{E}(X^s).
\]

(b) By converting \( Y^z \) into ordinary powers \( (3) \) the sequence of ordinary power moments \( \left(\mathbb{E}(Y^s)\right)_{s \in \mathbb{N}} \) 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:

\[
\mathbb{E}(Y^s) = \mathbb{E} \left( \sum_{j=0}^{s} \binom{s}{j} Y^j \right) = \sum_{j=0}^{s} \binom{s}{j} \mathbb{E}(Y^j) = \sum_{j=0}^{s} \binom{s}{j} \rho^j \mathbb{E}(X^j), \quad s \geq 1.
\]

The result concerning the moment generating function in (c) can be shown similar to (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 \( \psi\left(\rho(e^z - 1)\right) \), where \( \psi(z) = \mathbb{E}(e^{zX}) \) denotes the moment generating function of the mixing distribution \( X \). If the cdf of \( X \) is not known, we can compute the moment generating function \( \varphi(z) \) of \( Y \) utilizing only the moments sequences:

\[
\varphi(z) = \mathbb{E}(e^{zY}) = \sum_{s \geq 0} \mathbb{E}(Y^s)^s s! = \sum_{s \geq 0} \sum_{j=0}^s \binom{s}{j} \rho^j \mu_j \frac{z^s}{s!} = \sum_{j \geq 0} \rho^j \mu_j \sum_{s \geq j} \binom{s}{j} \frac{z^s}{s!}.
\]

Using the bivariate generating function identity of the Stirling numbers of the second kind (see Wilf [69])

\[
\sum_{n \geq 0} \sum_{k \geq 0} \binom{n}{k} \frac{z^n}{n!} u^k = e^{u(e^z-1)}, \tag{11}
\]
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 \( k_X(z) = \log(\varphi(z)) \) and \( k_Y(z) = \log(\psi(z)) \). For a proof of part (d) we refer the reader to Johnson, Kotz and Kemp [36].

In the applied probability literature, see [37, 70], given \( Y \overset{d}{=} MPo(\rho X) \) it is usually assumed that the cumulative distribution function of the mixing distribution \( X \) of 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 \( E(X^s) = \mu_s \in \mathbb{R}^+, s \geq 1 \). Hence, it is beneficial to we express the probability mass function of a mixed Poisson distributed random variable solely in terms of the moments of \( X \), justifying (8). Note that for specific mixed Poisson distributions different simpler formulas may exist (compare with Corollay 2).

**Proposition 2.** Let \( X \) denote a random variable with moment sequence given by \((\mu_s)_{s \in \mathbb{N}}\) such that \( \psi(z) = E(e^{zX}) \) exists in a neighbourhood of zero, including the value \( z = -\rho \). A random variable \( Y \) with factorial moments given by \( E(Y^s) = \rho^s \mu_s \) has a mixed Poisson distribution \( Y \overset{d}{=} MPo(\rho X) \) 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

\[ \mathbb{P}\{Y = \ell\} = \frac{\rho^\ell}{\ell!} E(X^\ell e^{-\rho X}) = \frac{\rho^\ell}{\ell!} \left( D_z^n \psi(z) \right) \bigg|_{z=-\rho} = \frac{\rho^\ell}{\ell!} \left( D_z^n \left( \sum_{s \geq 0} z^s \mu_s \frac{\rho^s}{s!} \right) \right) \bigg|_{z=-\rho} = \frac{\rho^\ell}{\ell!} \sum_{s \geq \ell} (-\rho)^{s-\ell} \binom{s}{\ell} \frac{\mu_s}{(s-\ell)!} \mu_s \frac{\rho^s}{s!} = \sum_{s \geq \ell} (-1)^{s-\ell} \binom{s}{\ell} \mu_s \frac{\rho^s}{s!}, \]

(12)

Alternatively, the formula for the probability mass function can formally be obtained directly from the definition

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

Interchanging summation and integration leads to the stated result. □
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 [28] 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
\[
\mathbb{E}(X_n^s) = \lambda_n^s \cdot \mu_s \cdot (1 + o(1)), \quad s \geq 1,
\]
with \( \lambda_n \) 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 [48]: 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 by its moment sequence:
\[
\sum_{s \geq 1} (\mu_{2s})^{-\frac{1}{2s}} = \infty.
\]
Note that (14) is satisfied, whenever \( \mathbb{E}(e^{zX}) \) exists in a neighbourhood of zero. We obtain the following result concerning mixed Poisson distributions.

**Lemma 1** (Uniqueness of mixed Poisson distributions). *The moments of a mixed Poisson distributed random variable \( Y \equiv MPo(\rho X) \), 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 \( Y \) the following way
\[
\mathbb{E}(Y^s) = \sum_{\ell \geq 0} k^s \mathbb{P}\{Y = \ell\} = \sum_{\ell = 0}^{2s-1} k^s \mathbb{P}\{Y = \ell\} + \sum_{\ell \geq 2s} k^s \mathbb{P}\{Y = \ell\} \\
\leq (2s)^s \cdot 1 + \sum_{\ell \geq 2s} 2^s k^s \mathbb{P}\{Y = \ell\} \leq (2s)^s + 2^s \mathbb{E}(Y^2).
\]
This implies that
\[
\mathbb{E}(Y^s) \leq 2^s(s^s + \mathbb{E}(Y^2)) \leq 2^s(s^s + \rho^s \mu_s) \leq 4^s(1 + \rho)^s \cdot \max\{s^s, \mu_s\}.
\]
Consequently,

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

such that

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

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 integer \(0 < r < s\) we have

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

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

\[
\sum_{s \geq 1} \min\left\{ \frac{1}{s}, m_s \right\} = \infty,
\]

which immediately implies the required result; note that we omitted the additional factor \(\frac{1}{4}\) for the sake of simplicity. If \(m_s\) is bounded away from zero this is immediately true. Hence, we assume in the following 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 \(\sum_{s \geq 1} m_s = \infty\) the equation 15 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 \([a_\ell, b_\ell] := \{a_\ell, a_\ell + 1, \ldots, b_\ell\}\) 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 \(\sum_{s \in I_2} \frac{1}{s}\) is finite and satisfies

\[
\sum_{s \in I_2} \frac{1}{s} = \sum_{\ell \in \mathbb{N}} \sum_{s \in [a_\ell, b_\ell]} \frac{1}{s} > \sum_{\ell \in \mathbb{N}} \int_{a_\ell}^{b_\ell} \frac{1}{x} \, dx = \sum_{\ell \in \mathbb{N}} \ln \left( \frac{b_\ell}{a_\ell} \right) > 0.
\]
Furthermore, \( \ln \left( \frac{b_\ell}{a_\ell} \right) < \epsilon \) for all sufficiently large \( \ell \) such that \( b_\ell < e^\epsilon a_\ell \). This implies that for \( k \in [a_\ell, b_\ell] \) and sufficiently large \( \ell \)
\[
\frac{m_k}{k} \leq \frac{m_{a_\ell}}{b_\ell} \leq \frac{m_{a_\ell}}{e^\epsilon a_\ell} \leq \frac{m_{a_\ell - 1}}{e^\epsilon a_\ell} \leq \frac{1}{e^\epsilon a_\ell} \leq 2\epsilon.
\]
Hence, \( m_k \leq \frac{2\epsilon}{k} \). 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 \( \sum_{s \in I_2} m_s = \infty \).

Concerning random discrete structure one usually encounters discrete distributions, which are supported on (a subset of) the non-negative integers. It is convenient to use factorial moments instead of the power moments, since they can be directly obtained from the probability generating function by repeated differentiation, see (7). Mixed Poisson distributions and a related phase transition naturally occur if the factorial moments satisfy asymptotic expansions similar to (13) instead of the power 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 \( \lambda_n > 0 \). Assume that and 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 \( \lambda_n \to \infty \) for \( n \to \infty \) the random variable \( \frac{X_n}{\lambda_n} \) converges in distribution, with convergence of all moments, to \( X \).

(ii) if \( \lambda_n \to \rho \in (0, \infty) \) for \( n \to \infty \) the random variable \( X_n \) converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable \( Y \equiv \text{MPo}(\rho X) \).

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

**Remark 1.** It may be possible to unify cases (i) and (ii) to arbitrary sequences \( \lambda_n \) 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 \( \lambda_n \). Thus, the distribution of \( \frac{X_n}{\lambda_n} \) degenerates for \( \lambda_n \to \infty \), since we expect a central limit theorem for \( (X_n - \lambda_n)/\sqrt{\lambda_n} \). It might also be necessary for non-generate \( X \) 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 \( \lambda_n \to 0 \), such that \( X_n \overset{\mathcal{L}}{\to} 0 \). It suffices to prove that \( \mathbb{E}(X_n) \to 0 \) and \( \mathbb{E}(X_n^2) \to 0 \). The general condition on the moments takes the form \( \mathbb{E}(X_n^s) = \lambda_n^s \cdot \mu_s \cdot (1 + o(\frac{1}{\lambda_n})) \).
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 \( n \to \infty \) 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 [42, 44, 57, 47] and also [32, 16]) is the following dichotomy for the asymptotic expansion of the factorial moments:
\[
\mathbb{E}(X_s^n) = \begin{cases} 
\lambda_{s,0} \cdot \mu_{s,j} \cdot (1 + o(1)), & s \geq 1, \quad j \text{ fixed}, \\
\lambda_{s,1} \cdot \mu_s \cdot (1 + o(1)), & s \geq 1, \quad j \to \infty,
\end{cases}
\]
where \( \lambda_{s,0} \) is independent of \( j \), but \( \lambda_{s,1} = \lambda_{s,1}(j) \) 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 additional property holds for \( s \geq 1 \):
\[
\Lambda_j \mu_{s,j} \to \mu_s, \quad j \to \infty,
\]
where \( \Lambda_j \) denotes an additional scale parameter; compare with the Remarks 6, 7, and 13.

Proof. By (3) the power moments of \( X_n \) satisfy the following asymptotic expansion
\[
\mathbb{E}(X_s^n) = \sum_{j=0}^{s} \binom{s}{j} \mathbb{E}(X_{s,j}^n) = \sum_{j=0}^{s} \binom{s}{j} \lambda_{s,j} \mu_{s,j} \cdot (1 + o(1)) = \left( \sum_{j=0}^{s} \binom{s}{j} \lambda_{s,j} \mu_{s,j} \right) (1 + o(1)).
\]
If \( \lambda_n \to \infty \) for \( n \to \infty \), we obtain further the expansion
\[
\mathbb{E}(X_s^n) = \left( \sum_{j=0}^{s} \binom{s}{j} \lambda_{s,j} \mu_{s,j} \right) (1 + o(1)) = \left( \sum_{j=0}^{s} \binom{s}{j} \lambda_{s,j} \mu_{s,j} + O(\lambda_{s,-1}^n) \right) (1 + o(1))
\]
\[
= \lambda_{s,0} \mu_s + O(\lambda_{s,-1}^n) + o(\lambda_{s,0}^n).
\]
Consequently, the moments of \( \frac{X_n}{\lambda_n} \) convergence to the moments \( \mu_s \) of the mixing distribution. By the Fréchet-Shohat moment convergence theorem and the moments of \( X \) satisfying Carleman’s condition, this proves convergence in distribution. Furthermore, for \( \lambda_n \to \rho \) for \( n \to \infty \), we directly obtain
\[
\mathbb{E}(X_s^n) = \left( \sum_{j=0}^{s} \binom{s}{j} \lambda_{s,j} \mu_{s,j} \right) (1 + o(1)) = \sum_{j=0}^{s} \binom{s}{j} \rho^j \mu_j + o(1).
\]
Consequently, the moments of $X_n$ converge to the moments of a mixed Poisson distributed random variable $Y \overset{d}{=} \text{MPo}(\rho X)$, which is uniquely determined by its moments 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 $Y \overset{d}{=} \text{MPo}(\rho X)$ converges to its mixing distribution for $\rho \to \infty$. □

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 $1 \leq j \leq n [42, 44, 45, 56]$. In the limit $n \to \infty$, with $j = j(n)$, phase transitions where observed according to the relative growth of $j$ with respect to $n$, i.e. $j = 1, 2, \ldots$ being a constant independent of $n$, $j \to \infty$ but with $j = o(n)$, or $j \sim \rho \cdot n$, for fixed $\rho$. As mentioned in the introduction, we will unify and simplify earlier arguments, starting from explicit formulas for the factorial moments from the various works. These explicit formulas directly lead to mixed Poisson laws, using Lemmas 1 and 2, and Stirling’s formula for the Gamma function

$$\Gamma(x) = x^{x-\frac{1}{2}}e^{-x} \sqrt{2\pi} \left(1 + O\left(\frac{1}{x}\right)\right) \quad \text{for} \quad x \to \infty. \hspace{1cm} (16)$$

Besides, whenever possible we interpret the random variables in terms of urn models.

3.1. Block sizes in $k$-Stirling permutations. Stirling permutations were defined by Gessel and Stanley [24]. A Stirling permutation is a permutation of the multiset $\{1, 1, 2, 2, \ldots, n, n\}$ 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., $1122, 1221, 2211$ 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 [24] for details, and [39] for a bijections with certain tree families. A straightforward generalization of Stirling permutations is to consider permutations of a more general multiset $\{1^k, 2^k, \ldots, n^k\}$, with $k \in \mathbb{N}$ (we use in this context $j^\ell := j, \ldots, j$, 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 permutations called $k$-Stirling permutations have already been considered by Brenti [8, 9], Park [59, 60, 61] and Janson [33, 35]. These $k$-Stirling permutations can be generated in a sequential manner: we start with $1^k = 11$ and insert the string $(n+1)^k$ at any position (anywhere, including first or last) in a given $k$-Stirling permutation of $\{1^k, 2^k, \ldots, n^k\}$, $n \geq 1$. 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 $\sigma = \sigma_1 \cdots \sigma_s$ is a substring $\sigma_a \cdots \sigma_b$, 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 $j \in \{1, 2, \ldots, n\}$, 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 $\sigma = 112233321445554777666$, has block decomposition $[112233321][445554][777][666]$.

The number of blocks of size $k \cdot \ell$ in a random $k$-Stirling permutation of order $n$ was studied in [45]. There, an simple exact expression for the factorial moments was derived:

$$E\left( X_{n,\ell}^s \right) = s! \left( \frac{\ell - 1 - \frac{1}{k}}{\ell - 1} \right)^s \cdot \frac{(n-\ell s + \frac{s^2}{2} - 1)}{(n-1 + \frac{1}{\ell})}. $$

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

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

(17)

could be characterized using observations by Janson et al. [35], and Janson [32]. It has a density function $f(x)$ that can be written as

$$f(x) = \frac{\Gamma\left(\frac{1}{k}\right)}{\pi} \sum_{j=1}^{\infty} (-1)^{j-1} \frac{\Gamma\left(\frac{j}{k} + 1\right) \sin \frac{j\pi}{k}}{j!} x^j, \quad x > 0; $$

(18)

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

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

(19)

Using Lemma 2, we can fill this gap, extending the results of [45].

**Corollary 1.** The factorial moments of 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 $n \to \infty$ asymptotically of mixed Poisson type, with mixing distribution $X$, determined by its moments and density given by (17) and (18) and scale parameter $\lambda_{n,\ell} = \frac{1}{k^2} \left( \ell - 1 \right) n^{\frac{1}{\ell}}$:

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

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

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

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

Moreover, for $\rho \to \infty$, the random variable $Y/\rho$ 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 $\alpha$ white balls and $\beta$ black balls, but if a black ball has been drawn, we put into the urn $\gamma$ white balls and $\delta$ black balls. The values $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ are fixed integer values and the urn model is specified by the $2 \times 2$ ball replacement matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. This definition readily extends to higher dimensions, leading to $r \times r$ ball replacement matrices. Note that one can we may consider $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, defining the urn process as a Markov process; see Remark 1.11 of Janson [32]. 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 [5, 16, 17, 29, 32, 63].

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

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

The initial configuration of the urn (it is here convenient to start here at time 1) is given by $(U_{1,0}, \ldots, U_{1,\ell+1}) = (2, k - 1, 0, \ldots, 0)$. It can be shown 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 [45], this implies that the random variables $U_{n,i}$ occurring in the urn model undergo a phase transition according to the growth of $\ell$ with respect to $n$, from continuous to discrete, where the moments of the occurring 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 \((-1 \ 0) \ 0 \ -1\). 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, 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 [47] and the references therein.

**Urn II.** Consider a possibly unbalanced generalized sampling without replacement urn model with ball replacement matrix 
\[
\begin{pmatrix}
-\alpha & 0 \\
0 & -\delta
\end{pmatrix}, \quad a, \delta \in \mathbb{N}.
\]
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, when all black balls have been drawn.

It was shown in [47] that the factorial moments of the random variable \(\hat{X}_{\delta m, \alpha n} = X_{\delta m, \alpha n}/\alpha\) are given by
\[
E\left(\hat{X}_{\delta m, \alpha n}^s\right) = \frac{n^s}{(m+\frac{\alpha s}{\delta})}, \quad s \geq 1.
\]
Moreover, a random variable \(Y\) arises in the limit, whose factorial moments are given by
\[
E(Y^s) = \rho s \Gamma(1 + \frac{\alpha s}{\delta}), \quad s \geq 1, \quad a, d \in \mathbb{R}.
\]
Using a special case of Theorem 2 it was shown that \(Y\) has a discrete distribution.

Assume that \(a/\delta \leq 2\):

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

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

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

**Remark 6.** As shown in [47], for fixed $m$ the random variable $\hat{X}_{B_m, \alpha n}/n$ converges to power of a beta-distributed random variable $B$, with moments $E(B^s) = 1/(m + \alpha s)$. The Weibull mixing distribution $X \sim W_{\delta/\alpha, 1}$ can be recovered by considering the limit $m \to \infty$ of $B = B_m$:

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

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 analyze the probability generating function $h_{n,m}(v)$, which can be derived using stochastic processes [46].

**Proof.** According to the definition of a mixed Poisson distributed random variable $Y \sim MPo(\rho X)$, it has factorial moments given by (20). 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

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

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

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

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

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

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

### 3.3. Descendants in Increasing trees

Increasing trees are labeled trees where the nodes of a tree of size $n$ are labelled by distinct integers of the set $\{1, \ldots, n\}$ 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. [6], and can be combinatorially described 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 $\varphi(t) :=$
forms, as studied by [58]: where we used the abbreviations RECT

Note that changing \( \phi \) and the total weights \( \sum_{k \geq 0} \varphi_k t^k \). The simple family of increasing trees \( T \) associated with a degree-weight generating function \( \varphi(t) \), can be described by the formal recursive equation

\[
T = \{1\} \times \left( \varphi_0 \cdot \{\epsilon\} \cup \varphi_1 \cdot T \cup \varphi_2 \cdot T \ast T \cup \varphi_3 \cdot T \ast T \ast T \cup \cdots \right) = \{1\} \ast \varphi(T),
\]

where \( \{1\} \) denotes the node labelled by 1, \( \times \) the Cartesian product, \( \cup \) the disjoint union, \( \ast \) the partition product for labelled objects, and \( \varphi(T) \) the substituted structure (see e. g., the books [23], [19]). Note that the elements of \( T \) are increasing plane trees, and that a tree with (out-)degrees \( d_1, \ldots, d_n \) is given weight \( \prod_{i=1}^n \varphi_{d_i} \). A tree of order \( n \) is chosen randomly with probabilities proportional to the weights, leading to the random trees of size \( n \) from \( T \).

Let \( T_n \) be the total weight of all such trees of order \( n \). It follows from (21) that the exponential generating function \( T(z) := \sum_{n \geq 1} T_n z^n / n! \) of the total weights satisfies the autonomous first order differential equation

\[
T'(z) = \varphi(T(z)), \quad T(0) = 0.
\]

We consider tree families having degree-weights of one of the three following forms, as studied by [58]: where we used the abbreviations RECT for recursive trees, GPORT for generalized plane recursive trees, and \( d \)–INCT for \( d \)-ary increasing trees.

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

Consequently, by solving (22), we obtain exponential generating function \( T(z) \)

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

and the total weights \( T_n \),

\[
T_n = \varphi_0 c_1^{n-1} (n-1)! \left(\frac{n-1 + \frac{c_2}{c_1}}{n-1}\right).
\]

Note that changing \( \varphi_k \) to \( ab^k \varphi_k \) for some positive constants \( a \) and \( b \) will affect the weights of all trees of a given order \( n \) by the same factor \( a^c b^{n-1} \), which does not affect the distribution of a random tree from the family. Hence, when considering random trees from these three classes, \( \varphi_0 \) 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 \( \alpha := -1 - \frac{c_2}{c_1} > 0 \) for (random) generalized plane recursive trees, and \( d := \frac{c_1}{c_2} + 1 \in 2, 3, \ldots \) for (random) \( d \)-ary increasing trees.

As shown by Panholzer and Prodinger [58], random trees in the three classes of families given in (23) can be grown as 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 out-degree \(d(v)\)) of the already grown tree of order \(i\) with probabilities \(p(v)\) given by

\[
p(v) = \begin{cases} 
\frac{1}{i}, & \text{for } \text{RECT,} \\
\frac{d(v)+\alpha}{(\alpha+1)i-1}, & \text{with } \alpha := -1 - \frac{c_1}{c_2} > 0, \text{ for } \text{GPORT,} \\
\frac{d-d(v)}{(d-1)n+1}, & \text{for } d - \text{INCT.} 
\end{cases}
\]

Moreover, it has been shown [58] that there are only the three classes of simple families that can be grown in this way (for suitable \(p(v)\)).

Let \(D_{n,j}\) denote random variable the counting the number of descendants - 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 [42] this random variable has been studied for the three forehand mentioned tree families using a generating functions approach. In the following we collect, and somewhat simplify the 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 [42]:

\[
\mathbb{E}(\hat{D}_{n,j}^s) = s! \frac{(n-j)(s+c_2)}{(j+1+s)} 
\]

with \(c_1, c_2\) as given in (23). Hence, for \(n \to \infty\) and \(j = j(n) \to \infty\), the factorial moments of \(\hat{D}_{n,j}\) are of mixed Poisson type by Stirling’s formula for the Gamma function (16), 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 \(n \to \infty\) and \(j = j(n) \to \infty\), factorial moments of mixed Poisson type with a Gamma mixing distribution \(X \sim \gamma(1, 1 + \frac{c_2}{c_1})\), and scale parameter \(\lambda_{n,j} = \frac{n-j}{j}\):

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

(i) for \(\lambda_{n,j} \to \infty\) the random variable \(\frac{\hat{D}_{n,j} - 1}{\lambda_{n,j}}\) convergences in distribution, with convergence of all moments, to \(X\).
(ii) for $\lambda_{n,j} \to \rho \in (0, \infty)$ the random variable $\hat{D}_{n,j}$ convergences in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \overset{C}{=} \text{MPo}(\rho X)$, which has a negative binomial distribution.

Remark 7. Note that for fixed $n$ the random variable $\hat{D}_{n,j}/n \overset{L}{\to} Z$, where $Z \overset{C}{=} \beta(1 + \frac{c_2}{c_1}, j - 1)$ is asymptotically beta-distributed (see [42]). One readily recovers the mixing distribution $X$ from $Z$ by taking the limit $j \to \infty$, using a well known result for beta-distributed random variables:

$$j Z_j \to X, \quad \text{for } j \to \infty,$$

with convergence of all moments.

Remark 8. Panholzer and Seitz [57] 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 a specific node in labelled families of evolving $k$-tree models.

The parameter descendants of node $j$ can be modelled using urn models: we encounter classical Pólya urns with non-standard initial values, depending on the number of draws. Note that Mahmoud and Smythe [49] used a similar approach to study the descendants of node $j$ in recursive trees, for fixed $j$ compared to $n$.

Urn III (Descendants in Increasing trees - Pólya urn). Consider a Pólya urn with ball replacement matrix

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

and initial conditions

$$W_0 = \begin{cases}
1, & \\
\alpha, & \\
d, &
\end{cases} \quad B_0 = \begin{cases}
j - 1, & \text{RECT}, \\
(j - 1)(1 + \alpha), & \text{GPORT}, \\
d + (j - 2)(d - 1) - 1, & \text{d - INCT},
\end{cases}$$

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

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

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-oriented increasing trees.** Let $X_{n,j}$ denote the random variable counting the outdegree of node $j$ in a generalized plane-oriented recursive tree of size $n$.

It has been shown in [44] using a generating functions approach that the factorial moments of the random variable $X_{n,j}$ are given by

$$
\mathbb{E}(X_{n,j}^s) = \frac{\Gamma(s - 1 - \frac{c_2}{c_1})}{\Gamma(-1 - \frac{c_2}{c_1})} \sum_{k=0}^s \binom{s}{k} (-1)^k \frac{\Gamma(n - \frac{c_2}{c_1}(s - 1 - k)) \Gamma(j + \frac{c_2}{c_1})}{\Gamma(j - \frac{c_2}{c_1}(s - 1 - k)) \Gamma(n + \frac{c_2}{c_1})},
$$

for $j \geq 2$ with $c_1, c_2$ as given in (23) such that $\alpha = -1 - \frac{c_2}{c_1}$. Lemma 2 and an application of Stirling's formula for the Gamma function (16) leads to the following result.

**Corollary 4.** The random variable $X_{n,j}$, counting the out-degree of node $j$ in a random generalized plane-oriented increasing tree of size $n$, $1 \leq j \leq n$, has for $n \to \infty$ and $j = j(n) \to \infty$, falling factorial moments of mixed Poisson type with a Gamma mixing distribution $X \overset{\text{d}}{=} \gamma(1, \alpha)$, 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 $\lambda_{n,j} \to \infty$ the random variable $\frac{X_{n,j}}{\lambda_{n,j}}$ convergences in distribution, with convergence of all moments, to $X$.

(ii) for $\lambda_{n,j} \to \rho \in (0, \infty)$ the random variable $X_{n,j}$ convergences in distribution, with convergence of all moments, to a mixed Poisson distributed random variable $Y \overset{\text{d}}{=} \text{MPo}(\rho X)$, which has a negative binomial distribution.

**Remark 9.** The limit law for fixed $j$ has been determined in [44]: the random variable $X_{n,j}/n^{1/(\alpha+1)}$ convergences in distribution to a random variable $Z$ characterized by its moments with a given density. In the next section we will discuss this model in more detail.

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

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

for $1 \leq j \leq n$. The out-degree $X_{n,j}$ of node $j$ in a generalized plane-oriented increasing 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

$$X_{n,j} \overset{d}{=} W_{n-j} - \alpha.$$  

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-oriented recursive trees.

Let $X_{n,j,k}$ denote the random variable, which counts the number of size-$k$ branches (= subtrees) 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 $X_{n,j}$ counting the outdegree of node labelled $j$ by

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

![Figure 3](image.png)

**Figure 3.** A size 15 plane-oriented recursive trees 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. [67] for the case of the root node $j = 1$ and for the instance of random recursive trees: they derived the distribution of $X_{n,1,k}$ and a limit law for it. Further they stated results for joint distributions. The analysis was extended in [43] to increasing tree families generated by a natural growth process (see Subsection 3.3). In particular, for generalized plane-oriented recursive trees with parameter $\alpha$ the following result was obtained for the factorial moments of $X_{n,j,k}$:

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

In [43] only the case of fixed $k$ was considered. We can easily use Lemma 2 and Stirling’s formula for the Gamma function (16) to obtain the following result.

**Corollary 5.** The random variable $X_{n,j,k}$, counting the the number of size-$k$ branches attached to node $j$ in a random generalized plane-oriented increasing tree of size $n$ has for fixed $j$, $n \to \infty$ and $1 \leq k \leq n-j$, falling factorial moments of mixed Poisson type with mixing distribution $Z$ supported on $[0, \infty)$, and scale parameter
Remark 10. Thus, the density functions of $L_n$ results similar to our earlier findings for the ordinary outdegree $X_n$. Consider a balanced urn with balls of $k$ colors and let the random vector $(U_{n,1}, \ldots, U_{n,k+3})$ count the number of balls of each color at time $n$ with $\ell \times \ell$ ball replacement matrix $M$ given by
\[
M = \begin{pmatrix}
1+\alpha & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 1 & \alpha & \cdots & \cdots & \cdots & 0 \\
0 & 0 & -\alpha & 2\alpha + 1 & \cdots & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \cdots & \cdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & -((k-1)\alpha + k - 2) \\
0 & 0 & \cdots & 0 & 0 & 0 & -((k\alpha + k - 1) = (k+1)\alpha + k) \\
0 & 0 & \cdots & 0 & 0 & 0 & 1+\alpha \\
\end{pmatrix}.
\]
The initial configuration of the urn (it is here convenient to start here at time 0) is given by $(U_{0,0}, \ldots, U_{0,k+3}) = ((j-1)(\alpha + 1), \alpha, 0, \ldots, 0)$. The random variables $U_{n,i}$, with $3 \leq i \leq k + 2$, 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 increasing tree of size $n$, as follows:
\[U_{n-j,i+2} = X_{n,j,i} + i\alpha, \quad 1 \leq i \leq k.\]
Moreover, $U_{n-j,2}$ is related to the outdegree $X_{n,j}$ by $U_{n-j,2} = X_{n,j} + \alpha$.

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 $\theta$ is a discrete-time stochastic process, whose value at any positive-integer time $n$ is one of the $B_n$ partitions of the set $[n] = \{1, 2, 3, \ldots, n\}$ (see Pitman [62]). The parameters $\alpha$ and $\theta$ satisfy $0 < a < 1$ and $\theta > -a$. Here $B_n$ denote the Bell number counting the number of partitions of an $n$-element set $B_0 = B_1 = 1, B_2 = 2, B_3 = 5, etc.$\footnote{See sequence A000110 in OEIS.} One images 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 takes place and 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 \( t_i \), \( 1 \leq i \leq k \), \( 1 \leq k \leq n \) of sizes \(|t_i|\). At time \( n+1 \) the element \( n+1 \) is either added to one of the existing parts \( t_i \in T \) with probability
\[
P\{n + 1 < c \ t_i \} = \frac{|t_i| - a}{n + \theta}, \quad 1 \leq i \leq k,
\]
or added to the partition \( T \) as a new singleton block with probability
\[
P\{n + 1 < c \ |T| + 1 \} = \frac{\theta + k \cdot a}{n + \theta}.
\]
This model thus assigns a probability to any particular partition \( T \) of \([n]\). We are interested in the distribution of 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. in order to analyze the number of tables of a certain size we study instead a variant of the growth rule for generalized plane-oriented recursive trees. Combinatorially, we consider a family \( T_{\alpha,\beta} \) of generalized plane-oriented 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 non-root nodes in the tree \( \varphi(t) = \frac{1}{(1-t)^\alpha}, \alpha > 0 \), and the corresponding family \( T \)
\[
T_{\alpha,\beta} = \{1\} \times \vartheta(T), \quad T = \{1\} \times \varphi(T).
\]
(27)
The weight of the trees \( T \) in \( T_{\alpha,\beta} \) are defined as
\[
w(T) := \vartheta_{d(\text{root})} \prod_{v \in T \setminus \{\text{root}\}} \varphi_{d(v)},
\]
where \( d(v) \) denotes the out-degree of node \( v \). The generating functions \( T_{\alpha,\beta}(z) = \sum_{n \geq 1} T_{\alpha,\beta} \frac{z^n}{n!} \) and \( T(z) = \sum_{n \geq 1} T_n \frac{z^n}{n!} \) of trees in \( T_{\alpha,\beta} \) and \( T \) thus satisfies the differential equations
\[
T'_{\alpha,\beta}(z) = \vartheta(T(z)), \quad T'(z) = \varphi(T(z)).
\]
The growth process generating randomly a tree in \( T_{\alpha,\beta} \) can be described in the following way: 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 out-degree \( d(v) \) of the already grown tree with probabilities \( p(v) \), given for non-root nodes \( v \) by
\[
P\{n + 1 < c \ v\} = \frac{d(v) + \alpha}{\beta + (\alpha + 1)n},
\]
and for the root by
\[
P\{n + 1 < c \ \text{root}\} = \frac{d(\text{root}) + \beta}{\beta + (\alpha + 1)n}.
\]
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 $[n] = \{1, \ldots, n\}$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{A size 15 plane-oriented recursive trees 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.}
\end{figure}

**Proposition 3** (Chinese restaurant process and generalized plane-oriented recursive trees). A random partition of $\{1, \ldots, n\}$ generated by the Chinese restaurant process with parameters $a$ and $\theta$ can be equivalently generated by the growth process of generalized plane-oriented recursive trees $T_{\alpha, \beta}$ of size $n + 1$. The parameters $a, \theta$ and $\alpha, \beta > 0$ are related by

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

**Remark 11.** In the relation above $\theta$ cannot be negative since $\beta$ is assumed to be positive. The correspondence above 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 $\vartheta(t) = (1 - t)^{-\beta} = 1 + \beta t + \ldots$ 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:

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

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 $T_{\alpha, \beta}$ with labels $\{0, 1, \ldots, n\}$ has $k$ branches $t_i$, $1 \leq i \leq k$ of sizes $|t_i|$. By the considerations
of [58] at time \( n + 1 \) the element \( n + 1 \) is either attached to one of the existing non-root nodes \( v \) with probability
\[
\mathbb{P}\{n + 1 < v\} = \frac{d(v) + \alpha}{\beta + (\alpha + 1)n}
\]
or to the root of the tree with probability
\[
\mathbb{P}\{n + 1 < \text{root}\} = \frac{d(\text{root}) + \beta}{\beta + (\alpha + 1)n} = \frac{k + \beta}{\beta + (\alpha + 1)n} = \frac{\beta}{\alpha + 1} + \frac{k}{\alpha + 1}.
\]
Consequently, element \( n + 1 \) is attached to one of the branches \( t_i \in T \) with probability
\[
\mathbb{P}\{n + 1 < c_{t_i}\} = \sum_{v \in t_i} \mathbb{P}\{n + 1 < c_v\} = \sum_{v \in t_i} \frac{d(v) + \alpha}{\beta + (\alpha + 1)n} = \frac{|t_i| - 1 + |t_i|\alpha}{\beta + (\alpha + 1)n} = \frac{|t_i| - \frac{1}{\alpha + 1} + |t_i|\alpha}{\beta + (\alpha + 1)n}.
\]
Thus, setting \( a = \frac{1}{\alpha + 1} \) and \( \theta = \frac{\beta}{\alpha + 1} \) proves the stated result. \( \square \)

**Theorem 1.** The random variable \( C_{n,j} \) counting the number of parts of size \( j \) in a partition of \( \{1, \ldots, n\} \) generated by the Chinese restaurant process is distributed as \( X_{n+1,j} \), the number of branches of size \( j \) of the root in a size \( n + 1 \) generalized plane-oriented recursive tree \( T_{\alpha,\beta} \) with \( a = \frac{1}{\alpha + 1} \) and \( \theta = \frac{\beta}{\alpha + 1} \):
\[
C_{n,j} \overset{\mathcal{L}}{=} X_{n+1,j}.
\]
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 \( \lambda_{n,j} = \frac{n}{\alpha + 1} \left( \frac{1 - \frac{1}{\alpha + 1}}{\frac{1}{\alpha + 1}} \right) \):
\[
\mathbb{E}(X_{n+1,j}^s) = \lambda_{n,j}^s \cdot \frac{\Gamma(s + \beta)\Gamma(\frac{\beta}{\alpha + 1})}{\Gamma(\beta)\Gamma(\frac{\beta + s}{\alpha + 1})} (1 + o(1)).
\]
(i) for \( \lambda_{n,j} \to \infty \) the random variable \( \frac{X_{n+1,j}}{\lambda_{n,j}} \) converges in distribution, with convergence of all moments, to \( Z \).
(ii) for \( \lambda_{n,j} \to \rho \in (0, \infty) \) the random variable \( X_{n+1,j} \) converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable \( Y \overset{\mathcal{L}}{=} \text{MPo}(\rho Z) \).

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

**Proof of Theorem 1.** We can study \( X_{n+1,j} \) the number of branches of size \( j \) attached to the root of a size \( n + 1 \) using the variable \( v \) as a marker and the generating function
\[
T_{\alpha,\beta}(z, v) = \sum_{n \geq 1} T_{\alpha,\beta,n}(z,v) \frac{z^n}{n!}.
\]
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:
\[ \mathbb{E}(X_{n+1,j}^s) = s! \frac{\left[z^n w^s\right] T'_\alpha,\beta(z,v)}{T_{\alpha,\beta,n}}, \]
where \( w = v - 1 \). Consequently, expansion of \( T'_\alpha,\beta(z,v) \) at \( w = 0 \) gives
\[ \mathbb{E}(X_{n+1,j}^s) = \frac{n! s!}{T_{\alpha,\beta,n}} \cdot \frac{(\beta - 1 + s) \left[z^n j^s\right]}{(1 - T(z))^{\beta + s}}. \]

Since \( T_{\alpha,\beta,n} = (\alpha + 1)^n \left(\frac{\beta + s + n - 1}{\alpha + 1}\right) \), we obtain the explicit result
\[ \mathbb{E}(X_{n+1,j}^s) = \frac{(\beta - 1 + s) \left[z^n j^s\right]}{(1 - T(z))^{\beta + s}}. \]

Application of Stirling’s formula for the Gamma function then leads to the stated result. \( \square \)

4. TRIANGULAR URN MODELS

In the study of node degree in generalized plane-oriented recursive trees 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
\[ \left( \begin{array}{c} \alpha \\ 0 \\ \gamma \end{array} \right), \quad \alpha, \beta, \gamma = \alpha + \beta \in \mathbb{N}. \]

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 after \( n \) draws.

This urn model has been studied by Puyhaubert [65, 16] who derived the probability mass function of \( W_n \), and a limit law for \( n \to \infty \). The results of [65, 16] were extended by Janson [32] 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 \( \sigma \)-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

$$\mathbb{E}(X_n \mid F_{n-1}) = \mathbb{E}(X_{n-1} + \Delta_n \mid F_{n-1}) = X_{n-1} + \mathbb{E}(\Delta_n \mid F_{n-1}).$$

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

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

Hence, let

$$X_n = \frac{X_n}{\alpha} = \frac{T_n}{T_{n-1}} = \frac{T_{n-1} + \alpha}{T_{n-1}} W_{n-1} \cdot \frac{T_{n-1}}{T_{n-1}} = X_{n-1}, \quad n \geq 1.$$

Then

$$\mathbb{E}(X_n \mid F_{n-1}) = X_n \cdot \frac{(n-1+\gamma T_0)}{(n-1+\gamma s+\gamma)} \cdot \frac{T_{n-1} + \alpha}{T_{n-1}} = X_{n-1}, \quad n \geq 1.$$

Consequently, $X_n$ is a positive martingale. By taking the unconditional expectation, this implies that the expected value of $X_n$ is given by

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

More generally, we similarly have for any positive integer $s$

$$\mathbb{E}\left(\left(\begin{array}{c} X_n + s - 1 \\ s \end{array}\right) \mid F_{n-1} \right) = \mathbb{E}\left(\left(\begin{array}{c} X_{n-1} + s - 1 \\ s \end{array}\right) \right) + \mathbb{E}\left(\left(\begin{array}{c} X_{n-1} + s - 1 \\ s \end{array}\right) \right) \frac{\alpha X_{n-1}}{T_{n-1}}$$

$$= \mathbb{E}\left(\left(\begin{array}{c} X_{n-1} + s - 1 \\ s \end{array}\right) \right) \frac{T_{n-1} + s\alpha}{T_{n-1}}.$$

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

$$\mathbb{E}\left(\left(\begin{array}{c} X_n + s - 1 \\ s \end{array}\right) \right) = \frac{(n-1+\gamma T_0)}{(n-1+\gamma s+\gamma)} \mathbb{E}\left(\left(\begin{array}{c} X_0 + s - 1 \\ s \end{array}\right) \right)$$

$$= \frac{(n+\gamma T_0)}{(n+\gamma (s+\gamma))} \frac{\Gamma(n+\gamma)}{\Gamma(n+\gamma+s)} \frac{\Gamma(\alpha)}{\Gamma(\alpha+s)} \frac{\Gamma(\beta)}{\Gamma(\beta+s)} \frac{\Gamma(\gamma)}{\Gamma(\gamma+s)} \frac{\Gamma(\gamma)}{\Gamma(\gamma+s)}.$$

**Theorem 2.** 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 $\left(\begin{array}{cc} \alpha & \beta \\ 0 & \gamma \end{array}\right)$, $\alpha, \beta, \gamma \in \mathbb{N}$, $\gamma = \alpha + \beta$, is given by the exact formula

$$\mathbb{E}(X_n^s) = \frac{\Gamma(n+\gamma T_0)}{\Gamma(n+\gamma (s+\gamma))} \frac{\Gamma(\alpha)}{\Gamma(\alpha+s)} \frac{\Gamma(\beta)}{\Gamma(\beta+s)} \frac{\Gamma(\gamma)}{\Gamma(\gamma+s)} \frac{\Gamma(\gamma)}{\Gamma(\gamma+s)}.$$

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

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

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

(ii) for \( \lambda_{n,b_0} \to \rho \in (0, \infty) \) the random variable \( \hat{X}_n \) convergences in distribution, with convergence of all moments, to \( Y \sim \text{MPo}(\rho X) \).

Remark 13. It is well known from the works of Puyhaubert [65, 16] and Janson [32], that for fixed \( b_0 \) the random variable \( X_n/n^\gamma \) tends to a random variable \( Z \) with moments

\[
E(Z_s) = \frac{\Gamma(w_0 + b_0) \Gamma(w_0 + s)}{\Gamma(w_0 + b_0 + s) \Gamma(w_0)} , \quad s \geq 1;
\]

for more details about the nature of this random variable we refer the reader to [16, 32]. 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 \( X \sim \gamma(w_0, 1) \) from \( Z \) as follows:

\[
\hat{Z}_{b_0} \xrightarrow{\rho} X_{b_0}, \quad \text{for} \ b_0 \to \infty,
\]

with convergence of all moments.

Proof. Let \( Y \) denote a random variable with rising factorial moments \( E(Y^\gamma) = E(Y(Y+1) \ldots (Y+s-1)) \) satisfying an expansion of mixed Poisson type, \( E(Y^\gamma) = \rho^s \cdot \mu_s \), for \( s \geq 1 \), with \( \mu_s \geq 0 \). We obtain the (falling) factorial moments using the binomial theorem for rising factorials (see [25]):

\[
x^\gamma = (x - s + 1)^\gamma = \sum_{\ell=1}^{s} \binom{s}{\ell} \cdot x^\gamma \cdot (-1)^{s-\ell} = \sum_{\ell=1}^{s} \binom{s}{\ell} \cdot x^\gamma \cdot (-1)^{s-\ell}, \quad s \geq 1.
\]

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

\[
(x + c)^\gamma = \sum_{j=0}^{\ell} \binom{\ell}{j} \cdot x^\gamma \cdot c^{\ell-j}, \quad \ell \geq 0.
\]

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
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 $\{1, 2, \ldots, |T|\}$, 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. $I_{n,i,j}$, 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 way.

$$\mathbb{E}(\hat{X}_n^r) = \sum_{\ell=1}^{s} \binom{s}{\ell} \mathbb{E}(\hat{X}_n^r)(-1)^{s-\ell}(s-1)^{s-\ell}$$

$$= \sum_{\ell=1}^{s} \binom{s}{\ell} (-1)^{s-\ell}(s-1)^{s-\ell} \sum_{j=0}^{\ell} (-1)^j \mathbb{E}(X_j^r)(-w_0/\alpha)^{\ell-j}.$$  

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

$$\mathbb{E}(X_n^s) = (\lambda_n b_0 + 1)^s \frac{\Gamma\left(\frac{w_0}{\alpha} + s\right)}{\Gamma\left(\frac{w_0}{\alpha}\right)} (1 + o(1)), \quad s \geq 1,$$

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

$$\mathbb{E}(\hat{X}_n^s) = s! \sum_{i=0}^{s} \lambda_{n,b_0}^i \left[ \sum_{j=i}^{s} \binom{j}{i} (-1)^j \left(\frac{w_0}{\alpha} + j - 1\right) \right] \times$$

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

Next, using the hypergeometric form of the Vandermonde convolution (see [25], p. 212), we obtain for the inner sum

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

We get further

$$s! \sum_{j=i}^{s} \binom{j}{i} \left(\frac{w_0}{\alpha} + j - s - \frac{w_0}{\alpha}\right) (-1)^j = \delta_{s,i} \cdot s! (-1)^s \frac{(-\frac{w_0}{\alpha})^s}{\frac{w_0}{\alpha}} = \delta_{s,i} \frac{\Gamma\left(\frac{w_0}{\alpha} + s\right)}{\Gamma\left(\frac{w_0}{\alpha}\right)} = \delta_{s,i} \frac{\Gamma\left(\frac{w_0}{\alpha} + s\right)}{\Gamma\left(\frac{w_0}{\alpha}\right)}.$$

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

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 $\{1, 2, \ldots, |T|\}$, 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. $I_{n,i,j}$, 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 way.
Figure 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.

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 [56] studied the r.v. \(I_{n,j}\) for random trees of so-called labelled simply generated tree families (see, e.g., [19]; 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 \(T\) 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) 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 out-degree of \(v\)). The family \(T\) 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 the total weight of all trees of size \(n\) in \(T\); 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 \(T\).

When analysing parameters in a simply generated tree family \(T\) 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 \(T\) 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 \(T\) of labelled simply generated trees and assuming the random tree model, in [56] the following asymptotic formula for the factorial moments of \(I_{n,j}\) has been obtained:

\[ E(I_{n,j}^s) = \Gamma \left( \frac{s}{2} + 1 \right) \left( \frac{2}{n} \right)^{\frac{s}{2}} \frac{(n-j)\xi}{n^2} \cdot (1 + o(1)). \]

where the constant \(\kappa\) depends on the particular tree family, i.e., on the degree-weight sequence, and is given in [56]. Consequently, an application of Lemma 2 and taking into account Example 3 yields the following result, which adds to the results of [56] the characterization of the limiting distribution as a mixed Poisson distribution.

\[ \textbf{Corollary 6.} \] 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
$n \to \infty$ 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

$$\lambda_{n,j} = \sqrt{\frac{1}{\kappa} \frac{n-j}{\sqrt{n}}} \quad \text{(with constant $\kappa$ given in [56])},$$

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

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

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

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

**Remark 14.** Considering $\lambda_{n,j}$, the critical phase occurs at $j = n - \Theta(\sqrt{n})$: $E(I_{n,j}) \to \infty$, for $n - j \gg \sqrt{n}$, whereas $E(I_{n,j}) \to 0$, 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 $\{r_1, \ldots, r_k\}$ 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 $\{S_1, \ldots, S_k\}$: $S_i$, $1 \leq i \leq k$, is the largest subtree rooted at the record $r_i$ not containing any of the remaining records $r_1, \ldots, r_{i-1}, r_{i+1}, \ldots, r_k$. 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 $\{1, 2, \ldots, |T|\}$ 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 $T$ of Cayley-trees can be described formally via the SET construction as

$$T = \circ \ast \text{SET}(T).$$

![Figure 6](https://example.com/figure6.png)
Note that Cayley-trees are a particular family of labelled simply generated trees as described in Subsection 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., [19, 66]) 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., a simply generated tree) has been studied by Janson [31] showing (after a suitable scaling by \(\frac{1}{\sqrt{n}}\)) a Rayleigh limiting distribution result; in particular, for Cayley-trees it holds \(\frac{R_n}{\sqrt{n}} \xrightarrow{d} \text{Rayleigh}(1)\). 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 [31], 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. \(C[v]_n\) 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., \(R_n \overset{d}{=} C[v]_n\). We can extend this relation by considering the r.v. \(C[v]_{n,j}\) 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 \(R_{n,j} \overset{d}{=} C[v]_{n,j}\). This can be seen easily by means of coupling arguments given in [31]: 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. It holds then 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 \(C[v]_{n,j}\) has factorial moments of mixed Poisson type yielding the following theorem.

**Theorem 3.** 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 \(n \to \infty\) 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 \(\lambda_{n,j} = \frac{\sqrt{n}j^{-1}}{j!e^j}\):  

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

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

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

**Remark 15.** Stirling’s formula for the Gamma function gives
\[
\lambda_{n,j} = \sqrt{\frac{n}{2\pi j^3}} \sim \sqrt{\frac{n}{2\pi j^3}}, \text{ for } j, n \to \infty, \text{ thus the critical phase occurs at } j = \Theta(n^{\frac{1}{3}}): \mathbb{E}(R_{n,j}) \to \infty,
\]
for \( j \ll n^{\frac{1}{3}} \), whereas \( \mathbb{E}(R_{n,j}) \to 0 \), 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.** We consider the description of the problem via the node-cutting procedure.
This immediately yields the following stochastic recurrence for the r.v. \( R_{n,j} \overset{\mathcal{L}}{=} C_{n,j} \):
\[
R_{n,j} \overset{\mathcal{L}}{=} R_{K_{n},j} + 1_{\{n-K_{n}=j\}}, \quad \text{for } 1 \leq j \leq n, \tag{29}
\]
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 (29) 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 \( \{1, \ldots, k\} \) and \( \{1, \ldots, n-k\} \), respectively, both subtrees are Cayley-trees of size \( k \) and \( n-k \), respectively. Consider now a particular pair \((T', T'')\) of Cayley-trees of sizes \( |T'| = k \) and \( |T''| = n-k \), 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 \((T', T'')\) of subtrees after cutting. By constructing such pairs \((T, x)\), one obtains that there are exactly \( k \binom{n}{k} \) possibilities \( (k \text{ possible ways of attaching } \hat{T}'' \text{ to a node in } \hat{T}' \text{ and } \binom{n}{k} \text{ possibilities of distributing the labels } \{1, \ldots, n\} \text{ order-preserving to the subtrees}) \), 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 \( p_{n,k} \):

\[
p_{n,k} = \mathbb{P}\{K_n = k\} = \begin{cases} \frac{kT_n}{nT_n}, & 1 \leq k \leq n - 1, \\ \frac{1}{n}, & k = 0, \end{cases}
\]

where we use throughout this section the abbreviation

\[
\tilde{T}_n := \frac{T_n}{n!} = \frac{n^{n-1}}{n!}.
\]

In order to treat the stochastic recurrence (29) 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 (29), 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:

\[
T(z) = \sum_{n \geq 1} \tilde{T}_n z^n = \sum_{n \geq 1} \frac{n^{n-1}}{n!} z^n.
\]

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. \( z \):

\[
F_z(z,v) = \frac{e^{T(z)} + (v - 1)T_j z^{j-1}}{1 - T(z) - (v - 1)T_j z^j}.
\]

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

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

We further obtain

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

It is not difficult to extract coefficients from (33) 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 \( \mathbb{E}(R_{n,j}^n) \). Namely, we use standard applications of so-called singularity analysis, see [19], to
transferring the local behavior of a generating function in a complex neighborhood 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 be-

havior of its coefficients. It holds (see, e.g., [19]) 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:

\[
T(z) = 1 - \sqrt{2} \sqrt{1 - ez} + \mathcal{O}(1 - ez).
\]

(34)

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

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

Singularity analysis yields then

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

Therefore, (33) yields

\[
[z^n w^s] \tilde{F}(z, w) = \mathcal{O}\left(\left(\frac{\tilde{T}_j}{e^j}\right)^s \epsilon^n n^{\frac{s+1}{2}}\right).
\]

This together with Stirling’s formula for the Gamma function (16) shows the follow-

ing bound for the \( s \)-th moments of \( R_{n,j} \), which holds uniformly for all \( 1 \leq j \leq n \):

\[
\mathbb{E}(R_{n,j}^s) = \frac{s!}{T_n} [z^n w^s] \tilde{F}(z, w) = \mathcal{O}\left(\left(\frac{\tilde{T}_j}{e^j}\right)^s \epsilon^n n^{\frac{s+1}{2}}\right) = \mathcal{O}\left(\left(\frac{\sqrt{n}}{e^{j}}\tilde{T}_j\right)^s\right). \tag{35}
\]

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

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

locally around \( z = e^{-1} \), which can be obtained from (34). Singularity analysis
gives then the expansion

\[
[z^n] \frac{1}{(1 - T(z))^{s+1}} = \frac{e^{n\frac{s+1}{2}}}{2^{s+1} (s+1)!} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right). \tag{36}
\]

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

\[
\mathbb{E}(R_{n,j}^s) = \frac{s!}{T_n} [z^n w^s] \tilde{F}(z, w) = \frac{s! \tilde{T}_j^s e^{n - js} (n - js)^{\frac{s+1}{2}}}{T_n 2^{s+1} (s+1)!} \cdot \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n - js}}\right)\right)
\]

\[
= \left(\frac{\sqrt{n}}{e^j}\right)^s \frac{s! \sqrt{\pi}}{2^{s+1} (s+1)!} \cdot \left(1 + o(1)\right) = \left(\frac{\sqrt{n}}{e^j}\right)^s 2^{s} \Gamma\left(\frac{s}{2} + 1\right) \cdot \left(1 + o(1)\right), \tag{37}
\]

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

\[
\frac{(s - 1)!}{2^{s-1} (\frac{s}{2} - 1)!} = \frac{\sqrt{\pi} (s - 1)!}{2^{s-1} (\frac{s}{2} - 1)!}. \tag{38}
\]
The mixed Poisson limit law with Rayleigh mixing distribution as stated in Theorem 3 follows then from (35) and (37) by applying Lemma 2.

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 Subsection 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 [50], 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 formulae 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 [54, 55] and in a more general setting by Janson [31]; in particular, for Cayley-trees one obtains $\frac{C_n}{\sqrt{n}} \xrightarrow{L} \text{Rayleigh}(1)$. Moreover, in [31] it was shown in general that for Galton-Watson tree families (thus containing Cayley-trees) the random variables $C_n$ and $C_n^{[e]}$ (as introduced in Subsection 5.2) for the edge-cutting procedure and the node-removal procedure, respectively, have the same limiting distribution behaviour. A number of works have analyzed the edge-cutting procedure and related processes using the connection between Cayley-trees and the so-called Continuum Random Tree, in particular see the work of Addagio-Berry, Broutin and Holmgren [1], and the recent works of Bertoin [3, 4].

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 $\{1, \ldots, |T| - 1\}$ 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 Subsection 5.2 (and stated already in [31]) one gets that the r.v. $R^{[e]}_n$ counting the number of edge-records in a random size-$n$ Cayley-tree is distributed as $C_n$, i.e., $R^{[e]}_n \overset{L}{=} C_n$. Moreover, the edge-records $e_1, \ldots, e_k$ of an edge-labelled tree $T$ naturally decompose $T$ into the root-node and $k$ edge-record subtrees $S_1, \ldots, S_k$, obtained from $T$ by removing the root-node of $T$ and all edges $e_1, \ldots, e_k$. Again we can introduce the r.v. $R^{[e]}_{n,j}$, 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 $R^{[e]}_{n,j} \overset{L}{=} C_{n,j}$.

In Figure 7 we illustrate the edge-cutting procedure for a particular tree.
The respective sizes of the subtrees cut-off during this procedure are 3, 2, and 1.

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 Subsection 5.2, but due to the less explicit nature of the formulæ occurring, the proof steps are more technical and a bit lengthy.

**Theorem 4.** 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 $n \to \infty$ 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

$$\lambda_{n,j} = \sqrt{n^{j-1}}.$$

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

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

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

**Remark 16.** According to Theorems 3 and 4 the r.v. $C_{n,j}$ and $C_{n,j}^e$ (and thus also $R_{n,j}^e$) and $R_{n,j}$, for the edge- and vertex-versions of the cutting procedures as considered in Subsection 5.2-5.3, have the same limiting distribution behaviour. JANSON [31] 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.** Decomposing a tree according to the first cut of the edge-cutting procedure immediately yields the following stochastic recurrence for the r.v. $C_{n,j}$:

$$C_{n,j} \overset{L}{=} C_{K_n^e,j} + 1_{\{n-K_n^e=j\}}, \text{ for } 1 \leq j < n,$$

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 [50] (and can be shown completely analogous to the computations in the proof of Theorem 3) that the random preservation property of Cayley-trees also holds for the edge-cutting
generating functions:

where we use the abbreviation

\[ T_n = \frac{n^{n-1}}{n!} \text{ given in (30)}: \]

\[ p_{n,k}^{[c]} := \mathbb{P}\{ K_n^{[c]} = k \} = \frac{kT_n T_{n-k}}{(n-1)T_n}, \quad \text{for } 1 \leq k \leq n-1 \text{ and } n \geq 2. \]

Again, in order to treat the stochastic recurrence (39) 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

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

where again the tree function \( T(z) = \sum_{n \geq 1} \tilde{T}_n z^n \) as defined in (31) appears. Solving this linear differential equation yields the following solution satisfying the initial condition \( G_j(0, v) = 0: \)

\[ G_j(z, v) = T(z) \cdot \exp \left( \int_0^z \frac{(v - 1) \tilde{T}_j t^{j-1}}{(1 - T(t))(1 - T(t) - (v - 1)\tilde{T}_j t^j)} \, dt \right). \quad (40) \]

To study the moments of \( C_{n,j} \) we apply in (40) 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. \( w: \)

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

where we use the abbreviation

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

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:

\[ G_j^{[s]}(z) := [w^s] \tilde{G}_j(z, w) = \sum_{n \geq 1} \frac{\tilde{T}_n \mathbb{E}(C_{n,j}^s)}{s!} z^n \]

\[ = T(z) \cdot \sum_{r=1}^s \frac{1}{r!} \sum_{\ell_1 + \cdots + \ell_q = s, \quad q=1 \atop \ell_q \geq 1, 1 \leq q \leq r} \prod_{q=1}^r \alpha_{\ell_q, j}(z), \quad \text{for } s \geq 1. \quad (41) \]

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 \( \ell \) arbitrary but fixed and uniformly for \( 1 \leq j \leq n \) and \( n \to \infty \):

\[
[z^n] \alpha_{j,\ell}(z) = \frac{T_j^\ell}{n} [z^{n-j\ell}] \frac{1}{(1-T(z))^{\ell+1}} = O\left(\frac{T_j^\ell e^{n \frac{\ell q_1-1}{2}}}{e^{\ell q_1}}\right).
\]

This implies

\[
[z^n] \prod_{q=1}^r \alpha_{j,\ell_q}(z) \bigg|_{\ell_1+\ldots+\ell_r=s, \ \ell_q \geq 1, 1 \leq q \leq r} = O\left(\frac{T_j^s e^{n \frac{s q_1-1}{2}}}{e^{s q_1}} \sum_{k_1+\ldots+k_r=n, \ k_q \geq 1, 1 \leq q \leq r} \prod_{q=1}^r \frac{\ell q_1-1}{2}\right). \tag{42}
\]

The sum occurring in the latter bound (42) can itself be bounded as follows:

\[
\sum_{k_1+\ldots+k_r=n, \ k_q \geq 1, 1 \leq q \leq r} \prod_{q=1}^r \frac{\ell q_1-1}{2} = \begin{cases} O\left(\frac{T_j^s e^{n \frac{s q_1-1}{2}}}{e^{s q_1}} \left(\log n\right)^Q\right), & \text{for } s > r, \\ O\left(\frac{(\log n)^{s-1}}{n}\right) , & \text{for } s = r, \end{cases}
\]

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

\[
[z^n] \prod_{q=1}^r \alpha_{j,\ell_q}(z) \bigg|_{\ell_1+\ldots+\ell_r=s, \ \ell_q \geq 1, 1 \leq q \leq r} = \begin{cases} O\left(\frac{T_j^s e^{n \frac{s q_1-1}{2}}}{e^{s q_1}} \left(\log n\right)^Q\right), & \text{for } s > r, \\ O\left(\frac{T_j^s e^{n (\log n)^{s-1}}}{e^{s q_1}}\right), & \text{for } s = r. \end{cases}
\]

Therefore, when considering the coefficients of the expression

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

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 \( n \to \infty \)):

\[
[z^n] \sum_{r=1}^s \frac{1}{r!} \sum_{\ell_1+\ldots+\ell_r=s, \ \ell_q \geq 1, 1 \leq q \leq r} \prod_{q=1}^r \alpha_{j,\ell_q}(z) = O\left(\frac{T_j^s e^{n \frac{s q_1-1}{2}}}{e^{s q_1}}\right). \tag{43}
\]

Now we are in a position to derive the stated bound on the \( s \)-th integer moments of \( C_{n,j} \). First, by using (41) and (43) we get for \( s \geq 1 \):

\[
E(C_{n,j}^s) = \frac{s!}{T_n} [z^n] G_j^s(z) = \frac{s!}{T_n} [z^n] T(z) \cdot \sum_{r=1}^s \frac{1}{r!} \sum_{\ell_1+\ldots+\ell_r=s, \ \ell_q \geq 1, 1 \leq q \leq r} \prod_{q=1}^r \alpha_{j,\ell_q}(z)
\]

\[
= O\left(\frac{s!}{T_n} \sum_{k=1}^{n-1} \frac{T_j^s e^{n-k \frac{k q_1-1}{2}}}{e^{s q_1}}\right) = O\left(\frac{T_j^s e^{n \frac{s q_1-1}{2}}}{e^{s q_1}} \sum_{k=1}^{n-1} \frac{(n-k)^\frac{s q_1-3}{2}}{k^{\frac{s}{2}}}\right).
\]

Splitting the remaining sum easily gives

\[
\sum_{k=1}^{n-1} \frac{(n-k)^\frac{s q_1-3}{2}}{k^{\frac{s}{2}}} = \sum_{k=1}^{n-1} \frac{(n-k)^\frac{s q_1-3}{2}}{k^{\frac{s}{2}}} + \sum_{k = \frac{n}{2}+1}^{n-1} \frac{(n-k)^\frac{s q_1-3}{2}}{k^{\frac{s}{2}}}.
\]
We split the summation interval of (48) at \( k \) of mixed Poisson type one has to spot and evaluate the main contribution of the remainder: 

\[ j, q \]

assume from now on that (36) and (44). The first part yields separately. Additionally we only require the already computed asymptotic bounds and furthermore (uniformly for such that

\[ 1 \leq j \leq n \]

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

\[ \mathbb{E}(C_{n,j}^s) = O\left( \left( \frac{\sqrt{n} \bar{T}_j}{e^j} \right)^s \right). \] (45)

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 (41) in more detail and to bound the remaining contributions. In order to do this we split

\[ \sum_{r=1}^{s} \frac{1}{j!} \prod_{q=1}^{r} \alpha_j, \ell_q(z) = \alpha_j, s(z) + \sum_{r=2}^{s} \prod_{q=1}^{r} \alpha_j, \ell_q(z), \]

such that

\[ \mathbb{E}(C_{n,j}^s) = \frac{s!}{T_n} \lfloor z^n \rfloor T(z) \cdot (\alpha_j, s(z) + Q_{j,s}(z)). \] (46)

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

\[ \lfloor z^n \rfloor Q_{j,s}(z) = O\left( \left( \frac{T_j T_s n^{\frac{s-2}{2}} \log n}{e^j s} \right)^s \right), \]

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

\[ \frac{s!}{T_n} \lfloor z^n \rfloor T(z) \cdot Q_{j,s}(z) = O\left( \left( \frac{\sqrt{n} \bar{T}_j}{e^j} \right)^s \cdot \frac{\log n}{\sqrt{n}} \right). \] (47)

Now we consider the term in (45) yielding the main contribution, where we assume from now on that \( j = o(n) \). We get

\[ \lfloor z^n \rfloor T(z) \alpha_j, s(z) = \sum_{k=1}^{n-s_j} \bar{T}_k \bar{T}_j^s \frac{1}{n-k} \lfloor z^{n-k-s_j} \rfloor \frac{1}{(1-T(z))^{s+1}}. \] (48)

We split the summation interval of (48) at \( k = \lfloor \frac{n}{s} \rfloor \) and consider the contributions separately. Additionally we only require the already computed asymptotic bounds (36) and (44). The first part yields

\[ \sum_{k=1}^{\lfloor \frac{n}{s} \rfloor} \bar{T}_k \bar{T}_j^s \frac{1}{n-k} \lfloor z^{n-k-s_j} \rfloor \frac{1}{(1-T(z))^{s+1}} \]

\[ = \bar{T}_j^s \sum_{k=1}^{\lfloor \frac{n}{s} \rfloor} \frac{1}{n-k} \frac{e^{n-k-s_j}(n-k-s_j)^{s-1}}{2^{s+1}(s-1)!} \cdot \left( 1 + O\left( \frac{1}{\sqrt{n} - k - s_j} \right) \right) \]

\[ = \frac{\bar{T}_j^s e^{n} n^{\frac{s-3}{2}}}{e^s 2^{s+1}(s-1)!} \sum_{k=1}^{\lfloor \frac{n}{s} \rfloor} \frac{1}{n-k} \left( 1 + O\left( \frac{k}{n} \right) + O\left( \frac{j}{n} \right) + O\left( \frac{1}{\sqrt{n}} \right) \right). \]
The main contribution comes from
\[
\sum_{k=1}^{\lfloor n/2 \rfloor} \frac{\tilde{T}_k}{e^k} = \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{\tilde{T}_k}{e^k} - \sum_{k=\lfloor n/2 \rfloor+1}^{\infty} \frac{\tilde{T}_k}{e^k} = \mathcal{T}(e^{-1}) + O\left( \sum_{k=\lfloor n/2 \rfloor+1}^{\infty} \frac{1}{k^s} \right) = 1 + O\left( \frac{1}{\sqrt{n}} \right),
\]
whereas
\[
O\left( \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{\tilde{T}_k}{e^k} \right) = O\left( \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k^s} \right) = O\left( \frac{1}{\sqrt{n}} \right).
\]
Thus
\[
\sum_{k=1}^{\lfloor n/2 \rfloor} \frac{\tilde{T}_k \tilde{T}_j}{n-k} \frac{1}{e^{n-k-sj}} \frac{1}{(1-T(z))^{s+1}} = \frac{\tilde{T}_j e^n n^{\frac{s-3}{2}}}{e^{sj} 2^{s+1} (s-1)!} \left( 1 + O\left( \frac{j}{n} \right) + O\left( \frac{1}{\sqrt{n}} \right) \right).
\]
(49)

The second part yields
\[
\sum_{k=\lfloor n/2 \rfloor+1}^{n-sj} \frac{\tilde{T}_k \tilde{T}_j}{n-k} \frac{1}{e^{n-k-sj}} \frac{1}{(1-T(z))^{s+1}} =
= O\left( \tilde{T}_j \sum_{k=\lfloor n/2 \rfloor+1}^{\lfloor n-sj \rfloor} \frac{e^{n-k-sj} (n-k-sj)^{\frac{s-1}{2}}}{(n-k)^{s+1}} \right) = O\left( \tilde{T}_j e^n \sum_{k=sj}^{\lfloor n/2 \rfloor-1} \frac{(n-sj)^{\frac{s-1}{2}}}{k} \right)
= O\left( \tilde{T}_j e^n \frac{\lfloor n/2 \rfloor-1}{e^{sj} n^{\frac{s-2}{2}}} \right), \quad \text{for } s \geq 2,
\]
\[
= O\left( \tilde{T}_j e^n \frac{e^{n \log n}}{e^{sj} n^{\frac{s}{2}}} \right), \quad \text{for } s = 1.
\]
(50)

Starting with (48) and combining the contributions (49) and (50) finally yields
\[
[z^n]T(z)\alpha_{j,s}(z) = \left( \sqrt{n} \tilde{T}_j \right)^s \frac{e^n}{2^{s+1} (s-1)! n^{\frac{s}{2}}} \left( 1 + O\left( \frac{j}{n} \right) + O\left( \frac{1}{\sqrt{n}} \right) \right),
\]
(51)

with \( \delta_{s,i} \) the Kronecker-delta function. Together with Stirling’s formula for the Gamma function we thus obtain from (46), (47) and (51)
\[
\mathbb{E}(C_{n,j}^s) = \frac{s!}{T_n} [z^n] C_{j}^{[s]}(z) = \frac{\sqrt{2\pi} s!}{2^{s+1} (s-1)!} \left( \sqrt{n} \tilde{T}_j \right)^s \left( 1 + O\left( \frac{j}{n} \right) + O\left( \frac{\log n}{\sqrt{n}} \right) \right).
\]

Finally, applying the duplication formula of the factorials (38), we get the stated result for the asymptotic expansion of the factorial moments:
\[
\mathbb{E}(C_{n,j}^s) = \left( \sqrt{n} \tilde{T}_j \right)^s 2^s \Gamma\left( \frac{s}{2} + 1 \right) \left( 1 + o(1) \right), \quad \text{for } j = o(n).
\]
(52)

The mixed Poisson limit law with Rayleigh mixing distribution as stated in Theorem 4 follows then from (45) and (52) by applying Lemma 2.

5.4. Parking functions and growth of the initial cluster. Parking functions are objects introduced by Konheim and Weiss [40], 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., [66], and computer science (e.g., due to close relations to hashing variants), see, e.g., [38]. 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 \( s_1, s_2, \ldots, s_n \). 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 \( s_k + 1, s_k + 2, \ldots \)) until 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 \((s_1, \ldots, s_n) \in \{1, \ldots, n\}^n\) such that all drivers are able to park. It has been shown already in [40] 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 \((s_1, \ldots, s_n) \in \{1, \ldots, n\}^n\), 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 \((s_1, \ldots, s_n)\) 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 Subsection 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 [66] 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

![Figure 8](image-url)
Figure 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.

has been used by Chassaing and Louchard in [11], where they related the growth 
of clusters in parking functions with the additive coalescent model for particle co-
agulation. Analogous to the definition in Subsection 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 \(x\) and \(x\).

**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 in-
crements 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 corre-
sponds to the number of max-record-subtrees of size \(j\) in the forests.

**Proof.** Given a parking function \((s_1, \ldots, s_n) \in \{1, \ldots, n\}^n\) 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 
\(\{1, \ldots, k\}\)). 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 
  spaces starting with parking space \(s + 1\) (i.e., the cluster of parking
Figure 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 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 \( \ell \)-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 \( \ell \) (by construction \( T' \not\in G \)). 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.

From the construction it follows 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 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.** 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., $X_n \overset{d}{=} C_n$. 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., $X_{n,j} \overset{d}{=} C_{n,j}$, $1 \leq j \leq n-1$.

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

$$
\frac{X_n}{\sqrt{n}} \overset{d}{\to} \text{Rayleigh}(1).
$$

$X_{n,j}$ has, for $n \to \infty$ 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 $\lambda_{n,j} = \frac{\sqrt{n}j^{j-1}}{j!}$:

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

(i) For $\lambda_{n,j} \to \infty$ the random variable $\frac{X_{n,j}^z}{\lambda_{n,j}}$ convergences in distribution, with convergence of all moments, to $X$. 
(ii) For \( \lambda_{n,j} \to \rho \in [0, \infty) \) the random variable \( X_{n,j} \) converges in distribution, with convergence of all moments, to a mixed Poisson distributed random variable \( Y \overset{d}{=} MPo(\rho X) \).

**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 \( \{1, \ldots, n\} \), 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

\[
X_n \overset{d}{=} C_n \quad \text{and} \quad X_{n,j} \overset{d}{=} C_{n,j}, \quad 1 \leq j \leq n-1.
\]

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 [31, 55] and Theorem 4, 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 [19][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.,

\[
\frac{X_n}{\sqrt{n}} \xrightarrow{d} \text{Rayleigh}(\sqrt{2}).
\]

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 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
be the combinatorial family of bridges of length \( \geq 0 \) and \( \mathcal{D} \) be the family of positive Dyck path excursions of length \( \geq 2 \), 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 \( \mathcal{D}^\bullet \) be the family of negative Dyck path excursions of length \( \geq 2 \), 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:

\[
\mathcal{B} = \text{SEQ}(\mathcal{D} \cup \mathcal{D}^\bullet).
\]  

(53)

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

\[
\mathcal{D} = \text{SEQ}(\mathcal{D}) \uparrow.
\]  

(54)

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 (53)-(54) 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 (53)-(54) 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

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

(55)

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

\[
\mathbb{E}(X_{n,j}^s) = \frac{s!}{B_n} [z^n w^s] \tilde{B}_j(z,w) = \frac{s!}{B_n} [z^n] \frac{2^s D_j^{s-j} z^{j s}}{(1 - 4z)^{\frac{j s}{2}}}
\]

\[
= \frac{s! 2^s D_j^{s-j} z^{j s}}{B_n [z^{n-j s}] (1 - 4z)^{\frac{j s}{2}}} \sim \frac{s! 2^s D_j^{s-j} (n-j s)^{\frac{j s}{2} - 1}}{B_n \Gamma\left(\frac{ss+1}{2}\right)}.
\]
If \( j = o(n) \) we further get
\[
\mathbb{E}(X_{n,j}^x) \sim \frac{s!2^s D_j^s 4^{n-j} s! n^{\frac{s-1}{2}}}{B_n \Gamma(s+1)\sqrt{n^s}},
\]
and, together with \( B_n \sim \frac{4^n}{\sqrt{\pi n^s}} \) and the duplication formula of the factorials (38), this gives
\[
\mathbb{E}(X_{n,j}^x) \sim \frac{s!2^s D_j^s 4^{n-j} s! \sqrt{n^s}}{\Gamma(s+1)\sqrt{n^s}} = 2^s D_j^s (4^{-j})^s n^{\frac{s-1}{2}} 2^s \Gamma\left(\frac{s}{2} + 1\right)
\]
Thus, by an application of Lemma 2 we get the following characterization of the limit law of \( X_{n,j} \).

**Theorem 6.** The random variable \( X_{n,j} \) counting the number of \( j \)-visits of the \( x \)-axes in a random bridge of length \( 2n \) has, for \( n \to \infty \) 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 \( \lambda_{n,j} = \frac{2\sqrt{2}D_j^s \Gamma\left(\frac{s}{2} + 1\right)}{4^{s/2}} \):
\[
\mathbb{E}(X_{n,j}^x) = (\lambda_{n,j})^s 2^s \Gamma\left(\frac{s}{2} + 1\right) \cdot (1 + o(1)).
\]

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

(ii) for \( \lambda_{n,j} \to \rho \in (0, \infty) \) the random variable \( X_{n,j} \) convergences in distribution, with convergence of all moments, to \( Y \overset{L}{\sim} \text{MPo}(\rho X) \).

Moreover, the random variable \( Y^x \overset{L}{\sim} \text{MPo}(\rho X) \) converges for \( \rho \to \infty \), after scaling, to its mixing distribution \( X \): \( Y^x \overset{L}{\to} X \), 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 [19]; see also Subsection 7.1.

### 5.6. Cyclic points and trees in graphs of random mappings.

We call a function \( f : [n] \to [n] \) from the finite set \( [n] := \{1, 2, \ldots, n\} \) 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., [13, 14, 18] 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 [13, 18]: 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
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 $\text{SET}$ and $\text{CYCLE}$ to the family of rooted labelled trees $\mathcal{T}$, as introduced and discussed in Subsection 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 $F(z) := \sum_{n \geq 0} n^n \frac{z^n}{n!} = \sum_{n \geq 0} \tilde{F}_n z^n$, with $\tilde{F}_n := \frac{n^n}{n!}$, of the number of $n$-mappings satisfies

$$F(z) = \exp \left( \log \left( \frac{1}{1 - T(z)} \right) \right) = \frac{1}{1 - T(z)}, \quad (56)$$

where the tree function $T(z)$ is defined in (31). Equation (56) 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 among 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 analyzed by Drmota and Soria [13]. They have shown a Rayleigh limit law: $X_n \overset{\mathcal{L}}{\to} \text{Rayleigh}(1)$. 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
Introducing vector components and cdf \( \Lambda \). Definition 2. Let a direct computation.

Moreover, the random variable \( Y \) converges for \( n \to \infty \) 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 \( \lambda_{n,j} \) equals \( \frac{j^{j-1} \sqrt{\pi}}{j e^j} \): 

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

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

(ii) for \( \lambda_{n,j} \to \rho \in (0, \infty) \) the random variable \( X_{n,j} \) converges in distribution, with convergence of all moments, to \( Y_{MPo}(\rho X) \).

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

6. Multivariate mixed Poisson distributions

The definition readily extends to multivariate distributions, compare with [15].

Definition 2. Let \((X_1, \ldots, X_m)\) denote a random vector with non-negative components and cdf \( \Lambda(\cdot) \) and \( \rho_1, \ldots, \rho_m > 0 \) scale parameters. The discrete random vector \((Y_1, \ldots, Y_m)\) with joint probability mass function given by 

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

\( \ell_1, \ldots, \ell_m \geq 0 \), has a multivariate mixed Poisson distribution with mixing distribution \((X_1, \ldots, X_m)\) and scale parameters \( \rho_1, \ldots, \rho_m \).

The relation for the moments extends to the multivariate case in the following way:

\[
\mathbb{E}(Y_{1}^{s_1} \cdots Y_{m}^{s_m}) = \rho_1^{s_1} \cdots \rho_m^{s_m} \mu_{s_1, \ldots, s_m}, \tag{57}
\]

where \( \mu_{s_1, \ldots, s_m} = \mathbb{E}(X_1^{s_1} \cdots X_m^{s_m}) \) for \( s_1, \ldots, s_m \geq 0 \); this can readily be seen by a direct computation.
Similar to Theorem 2 we obtain the following result when the distribution of the random vector \((X_1, \ldots, X_m)\) is uniquely determined by the sequence of its (mixed) moments.

**Proposition 5.** Let \(X = (X_1, \ldots, X_m)\) denote a random vector determined by its sequence of mixed moments \((\mu_{s_1, \ldots, s_m})_{s_1, \ldots, s_m \in \mathbb{N}_0}\) assuming that the moment generating function \(\psi(z) = \mathbb{E}(e^{zX})\) of \(X\) exists in a neighborhood of 0 including \(-\rho\). Then, the random vector \(Y = (Y_1, \ldots, Y_m)\) with mixed factorial moments given by (57) has a multivariate mixed Poisson distribution with mixing distribution \(X\) and scale parameters \(\rho_1, \ldots, \rho_m > 0\). The moment generating function \(\varphi(z) = \mathbb{E}(e^{zY})\) is given by the Stirling transform of \(\psi(z)\),

\[
\varphi(z) = \left. \left(\psi(z)\right)\right|_{z_1 = \rho_1(e^{z_1} - 1), \ldots, z_m = \rho_m(e^{z_m} - 1)}
\]

and the probability mass function of \(Y\) satisfies

\[
P\{Y_1 = \ell_1, \ldots, Y_m = \ell_m\} = \sum_{\ell_1 \geq j_1, \ldots, \ell_m \geq j_m} \mu_{j_1, \ldots, j_m} \prod_{i=1}^m \left(\frac{\rho_i}{\ell_i!} \cdot \frac{\rho_i^j}{j_i!}\right)^{\ell_i}
\]

for \(\ell_1, \ldots, \ell_m \geq 0\).

**Proof.** We proceed similarly to the proof of Theorem 2. Using

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

and (11), the moment generating function of \(Y\) is readily computed:

\[
\varphi(z) = \sum_{s_1, \ldots, s_m \geq 0} \mathbb{E}(Y_1^{s_1} \cdots Y_m^{s_m}) \frac{z_1^{s_1} \cdots z_m^{s_m}}{s_1! \cdots s_m!} = \sum_{\ell_1, \ldots, \ell_m \geq 0} \mu_{\ell_1, \ldots, \ell_m} \prod_{i=1}^m \left(\rho_i(e^{z_i} - 1)\right)^{\ell_i}
\]

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

\[
P\{Y_1 = \ell_1, \ldots, Y_m = \ell_m\} = \frac{\rho_1^{\ell_1} \cdots \rho_m^{\ell_m}}{\ell_1! \cdots \ell_m!} \cdot \left. \frac{\partial^{\sum_{k=1}^m \ell_k}}{\partial z_1^{\ell_1} \cdots \partial z_m^{\ell_m}} \varphi(z)\right|_{z=-\rho}
\]

and extension around 0. \(\square\)

Moreover, the basic limit theorem of Lemma 2 can be readily extended to limit laws for random vectors.

7. Outlook

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 Subsection 3.1 from [45], it is desirable to find a unifying scheme. The generating functions appearing indicate how to do so: we can use the critical compositions (see Flajolet and Sedgewick [19] Proposition IX.24, page 712), based on the pioneering works of Soria et al. [20, 21, 22, 14, 13, 2]. 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 substitution between structures of the form $F = G \circ H$. 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:

$$F(z, u, v) = G\left(u(H(z) - (v - 1)H_jz^j)\right).$$

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

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

such that $X_n = \sum_{j=1}^n X_{n,j}$. Using the semi-large power theorem of [19] (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 $(X_n; X_{n,j_1}, \ldots, X_{n,j_k})$ via

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

We will report on our findings on this refined analysis of compositions elsewhere [41].

7.2. **Open problems.** It should be possible to use mixed Poisson approximation and to derive distances, i.e. total variation distance - between the random variables of interest and the corresponding mixed Poisson distributions.

**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 authors warmly thank Walter Kuba for several motivating discussions and for providing a concise argument used in the series related part of Lemma 1.

**REFERENCES**

[1] L. Addario-Berry, N. Broutin, C. Holmgren, Cutting down trees with a Markov chainsaw. Preprint, arXiv:1110.6455, 2011.
[2] C. Banderier, P. Flajolet, G. Schaeffer, M. Soria, Random maps, coalescing saddles, singularity analysis and Airy phenomena, *Random Structures and Algorithms* Vol 19. 3-4, pp. 194–246, 2001.
[3] J. Bertoin, Fires on trees. *Annales de l’Institut Henri Poincaré Probabilités et Statistiques* 48, no. 4, 909–921, 2012.
[4] J. Bertoin and G. Miermont, The cut-tree of large Galton-Watson trees and the Brownian CRT. Preprint arXiv:1201.4081, 2012.
[5] Z.-D. Bai, F. Hu and L.-X. Zhang, Gaussian approximation theorems for urn models and their applications, *Annals of applied probability* 12, 4, 1149–1173, 2002.
[6] F. Bergeron, P. Flajolet and B. Salvy, Varieties of Increasing Trees, *Lecture Notes in Computer Science* 581, 24–48, 1992.
[7] M. Bernstein and N. J. A. Sloane, Some canonical sequences of integers. *Linear Algebra Appl.*, 226/228, 57–72, 1995.
[8] F. Brenti, Unimodal, log-concave, and Pólya frequency sequences in combinatorics, *Memoirs Amer. Math. Soc.*, 81, no. 413, 1989.
[9] F. Brenti, Hilbert polynomials in combinatorics, *J. Algebraic Combinatorics* 7, 127–156, 1998.
[10] T. Carleman, Sur le problème des moments, *Acad. Sci. Paris* 174, 1680–1682, 1922.
[11] P. Chassaing and G. Louchard, Phase transition for parking blocks, Brownian excursion and coalescence, *Random Structures & Algorithms*, 21, 76–119, 2002.
[12] J. H. Curtiss, A note on the theory of moment generating functions, *Annals of Math. Stat.* 13-4, 430–433, 1942.
[13] M. Drmota and M. Soria, Images and preimages in random mappings, *SIAM Journal on Discrete Mathematics*, 10, 246–269, 1997.
[14] M. Drmota and M. Soria, Marking in combinatorial constructions: Generating functions and limiting distributions, *Theoretical Computer Science*, 21-1, 76–119, 1999.
[15] A. Ferrari, G. Letacy and J.-Y. Tourneret, Multivariate mixed Poisson distributions, *EUSIPCO-04*, Vienna, Austria, 2004.
[16] 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.
[17] P. Flajolet, J. Gabarró and H. Pekari, Analytic urns, *Annals of Probability* 33, 1200–1233, 2005.
[18] P. Flajolet and A. Odlyzko, Random mapping statistics. in: *Advances in cryptology – EUROCRYPT’89*, Lecture Notes in Computer Science, 434, p. 329–354, Springer, Berlin, 1990.
[19] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*. Cambridge Univ. Press, Cambridge, UK, 2009.
[20] P. Flajolet and M. Soria, Gaussian limiting distributions for the number of components in combinatorial structures, *J. Combinatorial Theory, Series A*, vol 53:165–182, 1990.
[21] P. Flajolet and M. Soria, The cycle construction, *SIAM J. Discrete. Maths.*, vol.4(1): pp.58–60, February 1991.
[22] P. Flajolet and M. Soria, General combinatorial schemas: Gaussian limit distributions and exponential tails, *Discrete. Mathematics* vol.114: pp.159–180, 1993.
[23] P. Flajolet and J. Vitter, Average case analysis of algorithms and data structures, in *Handbook of Theoretical Computer Science*, 431–524, Elsevier, Amsterdam, 1990.
[24] I. Gessel and R. P. Stanley, Stirling polynomials, *J. Combin. Theory Ser. A* 24, no. 1, 24–33, 1978.
[25] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Second Edition. Reading, Massachusetts: Addison-Wesley, 1994.
[26] R. Grüber and N. Stefanoski, Mixed Poisson approximation of node depth distributions in random binary search trees. *Ann. Appl. Probab.*, Volume 15, 1A , 279–297, 2005.
[27] A. Gut, On the moment problem, *Bernoulli* Vol. 8, Nr. 3, 407–421, 2002.
[28] H.-K. Hwang and R. Neininger, Phase change of limit laws in the quicksort recurrence under varying toll functions. *SIAM Journal on Computing* 31, no. 6, 1687–1722, 2002.
[29] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns, *Stochastic processes and applications* 110, 177–245, 2004.
[30] S. Janson, Random records and cuttings in complete binary trees. In: *Mathematics and Computer Science III*, Algorithms, Trees, Combinatorics and Probabilities, M. Drmota, P. Flajolet, D. Gardy, B. Gittenberger (eds.), 241–253, Birkhäuser, Basel, 2004.
[31] S. Janson, Random cutting and records in deterministic and random trees. *Random Structures & Algorithms* 29, 139–179, 2006.
[32] S. Janson, Limit theorems for triangular urn schemes, *Probability Theory and Related Fields* 134, 417–452, 2005.
[33] S. Janson, Plane recursive trees, Stirling permutations and an urn model, Proceedings, Fifth Colloquium on Mathematics and Computer Science (Blaubeuren, 2008) , *DMTCS Proceedings A*, 541–548, 2008.
[34] S. Janson, Moments of Gamma type and the Brownian supremum process area. *Probability Surveys*, 7, 1–52, 2010.
[35] S. Janson, M. Kubá and A. Panholzer. Generalized Stirling permutations, families of increasing trees and urn models. *Journal of Combinatorial Theory Series A*, Volume 118, 94-114, 2011
[36] N. L. Johnson, S. Kotz and A. W. Kemp, *Univariate Discrete Distributions*, 2nd Edition, New York: John Wiley, 1992.

[37] D. Karlis and E. Xekalaki, Mixed Poisson Distributions, *International Statistical Review*, Vol. 73, No. 1, 35–58, 2005.

[38] D. E. Knuth, *The Art of Computer Programming, Volume 3: Sorting and Searching*, second edition, Addison-Wesley, Reading, 1998.

[39] L. M. Koganov, Universal bijection between Gessel-Stanley permutations and diagrams 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.

[40] A. G. Konheim and B. Weiss, An occupancy discipline and applications, *SIAM Journal on Applied Mathematics*, 14, 1266-1274, 1966.

[41] M. Kuba, A note on Critical Compositions and Mixed Poisson distributions, manuscript.

[42] M. Kuba and A. Panholzer, Descendants in increasing trees, *Electronic Journal of Combinatorics*, 13 (1) Paper 8, 2006.

[43] M. Kuba and A. Panholzer, Analysis of label-based parameters in increasing trees. *Discrete Mathematics and Theoretical Computer Science*, Proceedings AG, 321–330, 2006.

[44] M. Kuba and A. Panholzer, On the degree distribution of the nodes in increasing trees, *Journal of Combinatorial Theory, Series A* 114, 597–618, 2006.

[45] M. Kuba and A. Panholzer, Analysis of statistics for generalized Stirling permutations, *Combinatorics, Probability and Computing*, 20, 875–910, 2011.

[46] 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.

[47] M. Kuba and A. Panholzer, Limiting distributions for a class of diminishing urn models, *Advances in Applied Probability* 44, 1–31, 2012.

[48] M. Loève, *Probability Theory I*, 4th Edition, Springer-Verlag, New York, 1977.

[49] H. Mahmoud and R. Smythe, On the distribution of leaves in rooted subtrees of recursive trees, *Annals of Applied Probability* 1, 406–418, 1991.

[50] A. Meir and J. W. Moon, Cutting down random trees. *Journal of the Australian Mathematical Society* 11, 313–324, 1970.

[51] A. Meir and J. W. Moon, Cutting down recursive trees. *Mathematical Biosciences* 21, 173–181, 1974.

[52] J. C. Massé, and R. Theodorescu, Neyman type A distribution revisited. *Statistica Neerlandica*, Volume 59, Issue 2, 206-213, 2005.

[53] J. Neyman, On a new class of “contagious” distributions applicable in entomology and bacteriology, *Annals of Mathematical Statistics* 10, 35-57, 1939.

[54] A. Panholzer, Non-crossing trees revisited: cutting down and spanning subtrees. In: *Discrete Random Walks*, C. Banderier and C. Krattenthaler (eds.), *Discrete Mathematics and Theoretical Computer Science*, Proceedings AC, 265–276, 2003.

[55] A. Panholzer, Cutting down very simple trees. *Quaestiones Mathematicae* 29, 211–228, 2006.

[56] A. Panholzer and G. Seitz, Limiting distributions for the number of inversions in labelled tree families. Accepted for publication in *Annals of Combinatorics*.

[57] A. Panholzer and G. Seitz, Ancestors and descendants in evolving k-tree models. Accepted for publication in *Random Structures and Algorithms*.

[58] A. Panholzer and H. Prodinger, The level of nodes in increasing trees revisited, *Random Structures and Algorithms*, 31, 203–226, 2007.

[59] S. K. Park, The r-multipermutations, *J. Combin. Theory Ser. A* 67, no. 1, 44–71, 1994.

[60] S. K. Park, Inverse descents of r-multipermutations, *Discrete Math.* 132, no. 1–3, 215–229, 1994.

[61] S. K. Park, P-partitions and q-Stirling numbers, *J. Combin. Theory Ser. A* 68, no. 1, 33–52, 1994.

[62] J. Pitman, Exchangeable and partially exchangeable random partitions, *Probab. Theory Related Fields* 102, 145–158, 1995.

[63] N. Pouyanne, An algebraic approach to Pólya processes. *Annales de l’Institut Henri Poincaré*, Vol. 44, No. 2, 293–323, 2008.
[64] N. Privault, Generalized Bell polynomials and the combinatorics of Poisson central moments. Electronic Journal of Combinatorics 18, Paper 54, 2011.
[65] V. Puyhaubert, Modèles d’urnes et phénomènes de seuils en combinatoire analytique. Ph. D. thesis, École Polytechnique, Palaiseau, 2005.
[66] R. Stanley, Enumerative Combinatorics, Volume I & II, Cambridge University Press, 1997 & 1999.
[67] 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.
[68] Eric W. Weisstein, “Stirling Transform.” From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/StirlingTransform.html
[69] H. S. Wilf, Generatingfunctionology, Second Edition, Academic Press, 1992.
[70] G. Willmot, Mixed Compound Poisson Distributions, ASTIN, Volume 16, 59–80, 1986.

MARKUS KUBA, INSTITUTE OF APPLIED MATHEMATICS AND NATURAL SCIENCES, UNIVERSITY OF APPLIED SCIENCES - TECHNIKUM WIEN, HöCHSTÄDTPLATZ 5, 1200 WIEN
E-mail address: kuba@technikum-wien.at

ALOIS PANHOLZER, INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, TECHNISCHE UNIVERSITÄT WIEN, WIEDENER HAUPTSTR. 8-10/104, 1040 WIEN, AUSTRIA
E-mail address: Alois.Panholzer@tuwien.ac.at