Poisson-Kingman Partitions*

Jim Pitman

Technical Report No. 625

Department of Statistics
University of California
367 Evans Hall # 3860
Berkeley, CA 94720-3860

October 23, 2002

Abstract

This paper presents some general formulas for random partitions of a finite set derived by Kingman's model of random sampling from an interval partition generated by subintervals whose lengths are the points of a Poisson point process. These lengths can be also interpreted as the jumps of a subordinator, that is an increasing process with stationary independent increments. Examples include the two-parameter family of Poisson-Dirichlet models derived from the Poisson process of jumps of a stable subordinator. Applications are made to the random partition generated by the lengths of excursions of a Brownian motion or Brownian bridge conditioned on its local time at zero.

Keywords. exchangeable; stable; subordinator; Poisson-Dirichlet; distribution

1 Introduction

This paper presents some general formulas for random partitions of a finite set derived by Kingman's model of random sampling from an interval partition generated by subintervals whose lengths are the points of a Poisson point process. Instances

*Research supported in part by N.S.F. Grants MCS-9404345 and DMS-0071448

and variants of this model have found applications in the diverse fields of population genetics [17, 19], combinatorics [4, 48], Bayesian statistics [23], ecology [37, 15], statistical physics [11, 12, 13, 53, 55], and computer science [25].

Section 2 recalls some general results for partitions obtained by sampling from a random discrete distribution. These results are then applied in Section 3 to the Poisson-Kingman model. Section 4 discusses three basic operations on Poisson-Kingman models: scaling, exponential tilting, and deletion of classes. Section 5 then develops formulas for specific examples of Poisson-Kingman models. Section 6 recalls the two-parameter family of Poisson-Dirichlet models derived in [50] from the Poisson process of jumps of a stable(α) subordinator for $0 < \alpha < 1$. Section 7 reviews some results of [41, 46, 49, 50] relating the two-parameter family to the lengths of excursions of a Markov process whose zero set is the range of a stable subordinator of index α. Section 8 provides further detail in the case $\alpha = \frac{1}{2}$ which corresponds to partitioning a time interval by the lengths of excursions of a Brownian motion. As shown in [2, 3], it is this stable($\frac{1}{2}$) model which governs the asymptotic distribution of partitions derived in various ways from random forests, random mappings, and the additive coalescent. See also [5, 9] for further developments in terms of Brownian paths, and [10, 25] for applications to hashing and parking algorithms. This paper is a revision of the earlier preprint [42]. See [48] for a broader context and further developments.

2 Preliminaries

This section recalls some basic ideas from Kingman's theory of exchangeable random partitions [30, 31], as further developed in [43]. See [45, 48] for more extensive reviews of these ideas and their applications. Except where otherwise specified, all random variables are assumed to be defined on some background probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathbb{E}$ denotes expectation with respect to $\mathbb{P}$. Let $\mathbb{N} := \{1, 2, \ldots\}$, let F denote a random probability distribution on the line, and let Π be a random partition of $\mathbb{N}$ *generated by sampling from* F. That is to say, two positive integers i and j are in the same block of Π iff $X_i = X_j$, where conditionally given F the X_i are independent and identically distributed according to F. Formally, Π is identified with the sequence (Π_n), where Π_n is the restriction of Π to the finite set $\mathbb{N}_n := \{1, \ldots, n\}$. The distribution of Π_n is such that for each particular partition $\{A_1, \cdots, A_k\}$ of $\mathbb{N}_n$ with $\#(A_i) = n_i$ for $1 \leq i \leq k$, where $n_i \geq 1$ and $\sum_{i=1}^k n_i = n$,

$$\mathbb{P}(\Pi_n = \{A_1, \cdots, A_k\}) = p(n_1, \cdots, n_k) \tag{1}$$

for some symmetric function p of sequences of positive integers, called the *exchangeable partition probability function (EPPF)* of Π. Conversely, Kingman [30, 31] showed that if Π is an exchangeable random partition of $\mathbb{N}$, meaning that the distribution of its restrictions Π_n is of the form (1) for every n, for some symmetric function p, then Π has the same distribution as if generated by sampling from some random probability distribution F. Let P_i denote the size of the ith largest atom of F. If F is a random discrete distribution, then $\sum_j P_i = 1$ almost surely, and Π is said to have *proper frequencies* (P_i). In that case, let $\tilde{P}_j$ denote the size of the jth atom discovered in the process of random sampling. Put another way, $\tilde{P}_j$ is the asymptotic frequency of the jth class of Π when the classes are put in order of their least elements. It is assumed now for simplicity that $P_i > 0$ for all i almost surely, and hence $\tilde{P}_j > 0$ for all j almost surely. The sequence $(\tilde{P}_j)$ is a *size-biased permutation* of (P_i). That is to say, $\tilde{P}_j = P_{\pi_j}$ where for all finite sequences $(i_j, 1 \le j \le k)$ of distinct positive integers, the conditional probability of the event ($\pi_j = i_j$ for all $1 \le j \le k$) given $(P_1, P_2, \ldots)$ is

$$P_{i_1} \frac{P_{i_2}}{1 - P_{i_1}} \cdots \frac{P_{i_k}}{1 - P_{i_1} - \ldots - P_{i_{k-1}}}. \tag{2}$$

The distribution of Π_n is determined by the distribution of the sequence of ranked frequencies (P_i) through the distribution of the size-biased permutation $(\tilde{P}_j)$. To be precise, the EPPF p in (1) is given by the formula [43]

$$p(n_1, \cdots, n_k) = \mathbb{E}\left[\prod_{i=1}^{k} \tilde{P}_i^{n_i - 1} \prod_{i=1}^{k-1}\left(1 - \sum_{j=1}^{i} \tilde{P}_j\right)\right]. \tag{3}$$

Alternatively [45]

$$p(n_1, \cdots, n_k) = \sum_{(j_1, \ldots, j_k)} \mathbb{E} \prod_{i=1}^{k} P_{j_i}^{n_i} \tag{4}$$

where $(j_1, \ldots, j_k)$ ranges over all permutations of k positive integers, and the same formula holds with P_{j_i} replaced by $\tilde{P}_{j_i}$. For each $n = 1, 2, \cdots$ the EPPF p, when restricted to $(n_1, \cdots, n_k)$ with $\sum_i n_i = n$, determines the distribution of Π_n. Since Π_n is the restriction of Π_{n+1} to $\mathbb{N}_n$, the EPPF is subject to the following sequence of *addition rules* [43]: for $k = 1, 2, \ldots$

$$p(n_1, \cdots, n_k) = \sum_{j=1}^{k} p(\ldots, n_j + 1, \ldots) + p(n_1, \ldots, n_k, 1) \tag{5}$$

where $(\ldots, n_j + 1, \ldots)$ is derived from $(n_1, \ldots, n_k)$ by substituting $n_j + 1$ for n_j. The first few rules are

$$1 = p(1) = p(2) + p(1,1) \tag{6}$$

$$p(2) = p(3) + p(2,1); \qquad p(1,1) = 2p(2,1) + p(1,1,1) \tag{7}$$

where $p(2,1) = p(1,2)$ by symmetry of p. Let $\mu(q)$ denote the qth moment of $\tilde{P}_1$:

$$\mu(q) := \mathbb{E}[\,\tilde{P}_1^q\,] = \int_0^1 p^q \, \tilde{\nu}(dp). \tag{8}$$

where $\tilde{\nu}$ denotes the distribution of $\tilde{P}_1$ on $(0, 1]$. Following Engen [15], call $\tilde{\nu}$ the *structural distribution* associated with an random discrete distribution whose size-biased permutation is $(\tilde{P}_j)$, or with an exchangeable random partition Π whose sequence of class frequencies is $(\tilde{P}_j)$. The special case of (3) for $k = 1$ and $n_1 = n$ is

$$p(n) = \mathbb{E}[\,\tilde{P}_1^{n-1}\,] = \mu(n-1) \qquad (n = 1, 2, \cdots). \tag{9}$$

From (6), (7), and (9) the following values of the EPPF are also determined by the first two moments of the structural distribution:

$$p(1,1) = 1 - \mu(1); \;\; p(2,1) = \mu(1) - \mu(2); \;\; p(1,1,1) = 1 - 3\mu(1) + 2\mu(2). \tag{10}$$

So the distribution of the random partition of $\{1, 2, 3\}$ induced by Π with class frequencies $(\tilde{P}_i)$ is determined by the first two moments of the structural distribution of $\tilde{P}_1$. It is not true in general that the EPPF is determined for all $(n_1, \cdots, n_k)$ by the structural distribution, because it is possible to construct different distributions for a sequence of ranked frequencies which have the same structural distribution.

Continuing to suppose that (P_i) is the sequence of ranked atoms of a random discrete probability distribution, and that $(\tilde{P}_j)$ is a size-biased permutation of (P_i), for an arbitrary non-negative measurable function f, there is the well known formula

$$\mathbb{E}\left[\sum_i f(P_i)\right] = \mathbb{E}\left[\sum_j f(\tilde{P}_j)\right] = \mathbb{E}\left[\frac{f(\tilde{P}_1)}{\tilde{P}_1}\right] = \int_0^1 \frac{f(p)}{p}\tilde{\nu}(dp). \tag{11}$$

This formula shows that the structural distribution $\tilde{\nu}$ encodes much information about the entire sequence of random frequencies. Taking f in (11) to be the indicator of a subset B of $(0, 1]$, the quantity in (11) is $\nu(B) = \int_B p^{-1}\tilde{\nu}(dp)$. This measure ν is the mean intensity measure of the point process with a point at each $P_i \in (0, 1]$. For $x > \frac{1}{2}$ there can be at most one $P_i > x$, so the structural distribution $\tilde{\nu}$ determines the distribution of $P_1 = \max_j \tilde{P}_j$ on $(\frac{1}{2}, 1]$ via the formula

$$\mathbb{P}(P_1 > x) = \nu(x, 1] = \int_{(x,1]} p^{-1}\tilde{\nu}(dp) \qquad (x > \tfrac{1}{2}). \tag{12}$$

Typically, formulas for $\mathbb{P}(P_1 > x)$ get progressively more complicated on the intervals $(\frac{1}{3}, \frac{1}{2}]$, $(\frac{1}{4}, \frac{1}{3}]$, $\cdots$. See for instance [40, 50].

A random variable of interest in many applications is the sum of mth powers of frequencies

$$S_m := \sum_{i=1}^{\infty} P_i^m = \sum_{j=1}^{\infty} \tilde{P}_j^m \qquad (m = 1, 2, \dots)$$

where it is still assumed that $S_1 = 1$ almost surely. Let $\pi := \{A_1, \cdots, A_k\}$ be some particular partition of $\mathbb{N}_n$ with $\#(A_i) = n_i$ for $1 \leq i \leq k$, and consider the event

$(\Pi_n \geq \pi)$, meaning that each block of Π_n is some union of blocks of π. Then it is easily shown that

$$\mathbb{P}(\Pi_n \geq \pi) = \mathbb{E}\left[\prod_{i=1}^{k} S_{n_i}\right] = \sum_{j=1}^{k} \sum_{\{B_1,\ldots,B_j\}} p(n_{B_1},\ldots,n_{B_j}) \tag{13}$$

where the second sum is over partitions $\{B_1,\ldots,B_j\}$ of $\mathbb{N}_k$, and $n_B := \sum_{i\in B} n_i$. In particular, for $n_i \equiv m$ this gives an expression for the kth moment of S_m for each $k = 1, 2, \ldots$:

$$\mathbb{E}\left[S_m^k\right] = \sum_{j=1}^{k} \frac{1}{j!} \sum_{(k_1,\ldots,k_j)} \frac{k!}{k_1!\cdots k_j!}\, p(mk_1,\ldots,mk_j) \tag{14}$$

where the second sum is over all sequences of j positive integers $(k_1,\ldots,k_j)$ with $k_1 + \cdots + k_j = k$. Thus the EPPF associated with a random discrete distribution directly determines the positive integer moments of the power sums S_m, hence the distribution of S_m, for each m.

3 The Poisson-Kingman Model

Following McCloskey [37], Kingman [29], Engen [15], Perman-Pitman-Yor [40, 41, 50], consider the ranked random discrete distribution $(P_i) := (J_i/T)$ derived from an inhomogeneous Poisson point process of random lengths $J_1 \geq J_2 \geq \cdots \geq 0$ by normalizing these lengths by their sum $T := \sum_{i=1}^{\infty} J_i$. So it is assumed that the number N_I of J_i that fall in an interval I is a Poisson variable with mean $\Lambda(I)$, for some Lévy measure Λ on $(0,\infty)$, and the counts $N_{I_1},\cdots,N_{I_k}$ are independent for every finite collection of disjoint intervals $I_1,\cdots,I_k$. It is also assumed that

$$\int_0^1 x\Lambda(dx) < \infty \text{ and } \Lambda[1,\infty) < \infty$$

to ensure that $\mathbb{P}(T < \infty) = 1$. The sequence (P_i) may be regarded as a random element of the space $\mathcal{P}^{\downarrow}$ of decreasing sequences of positive real numbers with sum 1. Throughout this section, the following further assumption is made to ensure that various conditional probabilities can be defined without quibbling about null sets:

Regularity assumption. *The Lévy measure Λ has a density $\rho(x)$ such that the distribution of T is absolutely continuous with density*

$$f(t) := \mathbb{P}(T \in dt)/dt$$

which is strictly positive and continuous on $(0,\infty)$.

Note that the regularity assumption implies the total mass of the Lévy measure is infinite:

$$\int_0^\infty \rho(x)dx = \infty. \tag{15}$$

The results described below also have weaker forms for a Lévy density $\rho(x)$ just subject to (15), with appropriate caveats about almost everywhere defined conditional probabilities.

It is well known that f is uniquely determined by ρ via the Laplace transform

$$\mathbb{E}(e^{-\lambda T}) = \int_0^\infty e^{-\lambda x} f(x)dx = \exp[-\psi(\lambda)] \qquad (\lambda \geq 0) \tag{16}$$

where, according to the Lévy-Khintchine formula,

$$\psi(\lambda) = \int_0^\infty (1 - e^{-\lambda x})\rho(x)dx. \tag{17}$$

Alternatively, f is the unique solution of the following integral equation, which can be derived from (16) and (17) by differentiation with respect to λ:

$$f(t) = \int_0^t \rho(v) f(t-v) \frac{v}{t} dv. \tag{18}$$

Let $(\tilde{P}_j)$ be a size-biased permutation of the normalized lengths $(P_i) := (J_i/T)$ and let $(\tilde{J}_j) = (T\tilde{P}_j)$ be the corresponding size-biased permutation of the ranked lengths (J_i). Then (18) admits the following probabilistic interpretation [37, 41]:

$$\mathbb{P}(\tilde{J}_1 \in dv, T \in dt) = \rho(v) dv f(t-v) dt \frac{v}{t}. \tag{19}$$

This can be understood as follows. The left side of (19) is the probability that among the Poisson lengths there is some length in dv near v, and the sum of the rest of the lengths falls in an interval of length dt near $t - v$, and finally that the interval of length about v is the one picked by length-biased sampling. Formally, (19) is justified by the description of a Poisson process in terms of its Palm measures [41].

The following two Lemmas are read from [41, Theorem 2.1]. The first Lemma is immediate from (19), and the second is obtained by a similar Palm calculation.

Lemma 1 [41] *For each $t > 0$ the formula*

$$\tilde{f}(p \,|\, t) := pt\, \rho(pt) \frac{f(\bar{p}t)}{f(t)} \qquad (0 < p < 1;\ \bar{p} := 1 - p), \tag{20}$$

where ρ is the density of the Lévy measure of T and f is the probability density of T, defines a function of p which is a probability density on $(0,1)$. This is the density of the structural distribution of $\tilde{P}_1 := \tilde{J}_1/T$ given $T = t$:

$$\mathbb{P}(\tilde{P}_1 \in dp \,|\, t) = \tilde{f}(p \,|\, t) dp \qquad (0 < p < 1). \tag{21}$$

Lemma 2 [41] *For $j = 0, 1, 2, \cdots$ let*

$$T_j := T - \sum_{k=1}^{j} \tilde{J}_k = \sum_{k=j+1}^{\infty} \tilde{J}_k \tag{22}$$

which is the total length remaining after removal of the first j Poisson lengths $\tilde{J}_1, \ldots, \tilde{J}_j$ chosen by length-biased sampling. Then the family of densities (20) *on* $(0,1)$*, parameterized by $t > 0$, provides the conditional density of the random variable*

$$G_{j+1} := \frac{\tilde{J}_{j+1}}{T_j} = \frac{\tilde{P}_{j+1}}{\tilde{P}_{j+1} + \tilde{P}_{j+2} + \cdots}$$

given $T_0, \cdots, T_j$ via the formula

$$\mathbb{P}\left(G_{j+1} \in dp \,| T_0, \cdots, T_j\right) = \tilde{f}(p|T_j)\, dp \qquad (0 < p < 1). \tag{23}$$

Lemma 2 provides an explicit construction of a regular conditional distribution for $(\tilde{P}_j)$ given $T = t$ for arbitrary $t > 0$. This conditional distribution of $(\tilde{P}_j)$ given $T = t$ determines corresponding conditional distributions for the $\mathcal{P}^{\downarrow}$-valued ranked sequence (P_i) and for an associated random partition Π of $\mathbb{N}$.

Definition 3 The distribution of $(P_i) := (J_i/T)$ on $\mathcal{P}^{\downarrow}$ determined by the ranked points J_i of a Poisson process with Lévy density ρ will be called the *Poisson-Kingman distribution with Lévy density* ρ, and denoted $\textsc{pk}(\rho)$. Denote by $\textsc{pk}(\rho \,|\, t)$ the regular conditional distribution of (P_i) given $(T = t)$ constructed above. For a probability distribution γ on $(0, \infty)$, let

$$\textsc{pk}(\rho, \gamma) := \int_0^\infty \textsc{pk}(\rho \,|\, t)\gamma(dt) \tag{24}$$

be the distribution on $\mathcal{P}^{\downarrow}$ obtained by mixing the $\textsc{pk}(\rho \,|\, t)$ with respect to $\gamma(dt)$. Call $\textsc{pk}(\rho, \gamma)$ the *Poisson-Kingman distribution with Lévy density ρ and mixing distribution γ.*

Note that $\textsc{pk}(\rho \,|\, t) = \textsc{pk}(\rho, \delta_t)$, where δ_t is a unit mass at t, and that $\textsc{pk}(\rho) = \textsc{pk}(\rho, \gamma)$ for $\gamma(dt) = f(t)dt$. A formula for the joint density of $(P_1, \cdots, P_n)$ for (P_i) with $\textsc{pk}(\rho \,|\, t)$ distribution was obtained by Perman [40] in terms of the joint density $p_1(t, x)$ of T and J_1. This function can be described in terms of ρ and f as the solution of an integral equation [40], or as a series of repeated integrals [50]. But this formula will not be used here.

For a probability distribution Q on $\mathcal{P}^{\downarrow}$, such as $Q = \textsc{pk}(\rho, \gamma)$, a random partition Π of $\mathbb{N}$ will be called a *Q-partition* if Π is an exchangeable random partition of $\mathbb{N}$ whose ranked class frequencies are distributed according to Q. Immediately from

Definition 3, the structural distribution of a PK(ρ,γ)-partition Π of $\mathbb{N}$, that is the distribution on $(0,1)$ of the frequency $\tilde{P}_1$ of the class of Π containing 1, has density

$$\mathbb{P}(\tilde{P}_1 \in dp)/dp = \int_0^\infty \tilde{f}(p\,|\,t)\gamma(dt) \qquad (0 < p < 1) \tag{25}$$

where $\tilde{f}(p\,|\,t)$ given by (20) is the density of the structural distribution of $\tilde{P}_1$ given $T = t$ in the basic Poisson construction. Similarly, the EPPF of Π is

$$p(n_1, \cdots, n_k) = \int_0^\infty p(n_1, \cdots, n_k\,|\,t)\gamma(dt) \tag{26}$$

where $p(n_1, \cdots, n_k\,|\,t)$, the EPPF of a PK$(\rho\,|\,t)$-partition, is determined as follows:

Theorem 4 *The EPPF of a* PK$(\rho\,|\,t)$*-partition is given by the formula*

$$p(n_1, \cdots, n_k\,|\,t) = t^{k-1} \int_0^1 p^{n+k-2} I(n_1, \cdots, n_k; tp)\tilde{f}(p\,|\,t)dp \tag{27}$$

where $n = \sum_1^k n_i$, $I(n; v) = 1$ *if* $k = 1$ *and* $n_1 = n$, *and for* $k = 2, 3, \ldots$

$$I(n_1, \cdots, n_k; v) := \frac{1}{\rho(v)} \int_{\mathcal{S}_k} \left[\prod_{i=1}^k \rho(vu_i)u_i^{n_i}\right] du_1 \cdots du_{k-1} \tag{28}$$

where $\mathcal{S}_k$ *is the simplex* $\{(u_1, \ldots, u_k) : u_i \geq 0$ *and* $u_1 + \cdots + u_k = 1\}$.

Proof. In view of the formula (20) for $\tilde{f}(p\,|\,t)$, the formula (27) is obtained from formula (31) in the following Lemma by dividing by $f(t)dt$, letting $p = \sum_i x_i/t$, and integrating out with respect to p and to $u_i = x_i/(pt)$ for $1 \leq i \leq k-1$. □

A change of variables gives the following variant of formula (27), whose connection to the next lemma is a bit more obvious:

$$p(n_1, \cdots, n_k\,|\,t) = \int_0^t dv \frac{f(t-v)}{t^n f(t)} v^{n+k-1} I(n_1, \ldots, n_k; v)\rho(v). \tag{29}$$

Lemma 5 *Let* Π_n *be the restriction to* $\mathbb{N}_n$ *of a* PK(ρ) *partition* Π *whose class frequencies (in order of least elements) are* $\tilde{P}_j = \tilde{J}_j/T$, *where* $T = \sum_j \tilde{J}_j$ *has density* f, *and the lengths* $\tilde{J}_j$ *are the points of a Poisson process of lengths with intensity* ρ, *in length-biased random order. Then for each partition* $\{A_1, \cdots, A_k\}$ *of* $\mathbb{N}_n$ *such that* $\#(A_i) = n_i$ *for* $1 \leq i \leq k$,

$$\mathbb{P}(\Pi_n = \{A_1, \cdots, A_k\}, \tilde{J}_i \in dx_i \text{ for } 1 \leq i \leq k, T \in dt) \tag{30}$$

$$= t^{-n} f(t - \textstyle\sum_{i=1}^k x_i)\, dt \displaystyle\prod_{i=1}^k \rho(x_i)x_i^{n_i}\, dx_i. \tag{31}$$

Proof. This can be derived by evaluation of the expectation (3) for the joint distribution of $\tilde{P}_1, \ldots, \tilde{P}_k$ given $T = t$ determined by Lemma 2. Alternatively, there is the following more intuitive argument, which can be made rigorous using the characterization of Poisson process by its a Palm measures, as in [49, 41]. Let Π be constructed as in [46] using random intervals I_i laid down on $[0, T]$ in some arbitrary random order, where the lengths $J_i := |I_i|$ are the ranked points of the Poisson process with intensity $\rho(x)$, and $T = \sum_i J_i$. Let $U_1, U_2, \cdots$ be i.i.d. uniform on $(0, 1)$ independent of this construction. Let Π be the partition of $\mathbb{N}$ generated by the random equivalence relation $n \sim m$ iff either $n = m$ or TU_n and TU_m fall in the same interval I_i for some i. Then by construction, Π is a PK(ρ) partition. For the event in (30) to occur,

(i) there must be some Poisson point in dx_i for each $1 \le i \le k$, and

(ii) given (i), the sum of the rest of the Poisson points must fall in an interval of length dt near $t - \sum_{i=1}^k x_i$, and

(iii) given (i) and (ii), for each $1 \le i \le k$ and each $m \in A_i$ the sample point TU_m must fall in the interval of length x_i.

The infinitesimal probability in (30) therefore equals

$$\left(\prod_{i=1}^k \rho(x_i)\, dx_i \right) \; f(t - \textstyle\sum_{i=1}^k x_i)\, dt \displaystyle\prod_{i=1}^k \left(\frac{x_i}{t} \right)^{n_i} \tag{32}$$

which rearranges as (31). □

The formula (27) expresses $p(n_1, \cdots, n_k \,|\, t)$ as the expectation of a function of $\tilde{P}_1$ given $T = t$, where the function depends on t and $n_1, \cdots, n_k$. Because some values of an EPPF can always be expressed as moments of $\tilde{P}_1$, as in (8) and (10), it seems natural to try to express an EPPF similarly whenever possible. This idea serves as a guide to simplifying calculations in a number of particular cases treated later. The integrations in (27) and (28) are essentially convolutions, which can be expressed or evaluated in various ways. Consider for instance the length $T_k := T - \sum_{i=1}^k \tilde{J}_i$ which remains after removal of the first k lengths discovered by the sampling process. Then the formula of Lemma 5 can be recast as

$$\mathbb{P}(\Pi_n = \{A_1, \cdots, A_k\}, \tilde{J}_i \in dx_i \text{ for } 1 \le i \le k, T_k \in dv) \tag{33}$$

$$= (v + \textstyle\sum_{i=1}^k x_i)^{-n} \, f(v) dv \displaystyle\prod_{i=1}^k \rho(x_i) x_i^{n_i}\, dx_i \tag{34}$$

which yields the following integrated forms of (27):

Corollary 6 *The EPPF of a* PK(ρ)*-partition is given by the formula*

$$p(n_1, \cdots, n_k) = \int_0^\infty \cdots \int_0^\infty \frac{f(v) dv \prod_{i=1}^k \rho(x_i) x_i^{n_i}\, dx_i}{(v + \sum_{i=1}^k x_i)^n} \tag{35}$$

where $n := \sum_{i=1}^k n_i$*, or again by*

$$p(n_1, \cdots, n_k) = \frac{(-1)^{n-k}}{\Gamma(n)} \int_0^\infty \lambda^{n-1}\, d\lambda e^{-\psi(\lambda)} \prod_{i=1}^k \psi_{n_i}(\lambda) \tag{36}$$

where $\psi(\lambda) := \int_0^\infty (1-e^{-\lambda x})\rho(x)dx$ *is the Laplace exponent as in* (17)*, and*

$$\psi_m(\lambda) := \frac{d^m}{d\lambda^m}\psi(\lambda) = (-1)^{m-1}\int_0^\infty x^m e^{-\lambda x}\rho(x)dx \quad (m = 1, 2, \ldots). \tag{37}$$

Proof. Formula (34) yields (35) by integration, and (36) follows after applying the formula $b^{-n} = \Gamma(n)^{-1}\int_0^\infty \lambda^{n-1}e^{-\lambda b}d\lambda$ to $b = v + \sum_{i=1}^k x_i$. □

These integrated forms (35) and (36) also hold more generally, with $f(v)dv$ replaced by $\mathbb{P}(T \in dv)$, and $\rho(x)dx$ replaced by the corresponding Lévy measure on $(0,\infty)$, assuming only that the Lévy measure has infinite total mass.

Provided $\mathbb{E}(e^{\varepsilon T}) < \infty$ for some $\varepsilon > 0$, the Laplace exponent ψ can be expanded in a neighbourood of 0 as

$$\psi(\lambda) = -\sum_{m=1}^\infty \frac{\kappa_m}{m!}(-\lambda)^m$$

where the *cumulants* κ_m of T are the moments of the Lévy measure

$$\kappa_m = (-1)^{m-1}\psi_m(0) = \int_0^\infty x^m \rho(x)dx.$$

Then for each partition $\{A_1, \cdots, A_k\}$ of $\mathbb{N}_n$ such that $\#(A_i) = n_i$ for $1 \le i \le k$, Lemma 5 yields the formula

$$\mathbb{P}(\Pi_n = \{A_1, \cdots, A_k\}, T \in dt) = t^{-n}\mathbb{P}(T + \Sigma_{i=1}^k J_{i,n_i} \in dt)\prod_{i=1}^k \kappa_{n_i} \tag{38}$$

where J_{i,n_i} denotes a random length distributed according to the Lévy density tilted by x^{n_i}:

$$\mathbb{P}(J_{i,n_i} \in dx) = \kappa_{n_i}^{-1}\rho(x)x^{n_i}\, dx$$

and T and the J_{i,n_i} for $1 \le i \le k$ are assumed to be independent. If $f_{n_1,\ldots,n_k}(t)$ denotes the probability density of $T + \Sigma_{i=1}^k J_{i,n_i}$, then formula (27) for the EPPF of a $\text{PK}(\rho\,|\,t)$-partition can be rewritten

$$p(n_1, \cdots, n_k\,|\,t) = \frac{f_{n_1,\ldots,n_k}(t)}{t^n f(t)}\prod_{i=1}^k \kappa_{n_i} \tag{39}$$

and formula (35) for the EPPF of a $\text{PK}(\rho)$-partition becomes

$$p(n_1, \cdots, n_k) = \mathbb{E}\left[(T + \Sigma_{i=1}^k J_{i,n_i})^{-n}\right] \prod_{i=1}^{k} \kappa_{n_i}. \tag{40}$$

See also James [23] for closely related formulas, with applications to Bayesian non-parametric inference.

4 Operations

Later discussion of specific examples of Poisson-Kingman partitions will be guided by a number of basic operations on Lévy densities ρ and their associated families of partitions.

4.1 Scaling

By an obvious scaling argument, the $\text{PK}(\rho)$ and $\text{PK}(\rho')$ distributions are identical whenever $\rho'(x) = b\rho(bx)$ is a rescaling of ρ for some $b > 0$. The converse is less obvious, but true [49, Lemma 7.5].

4.2 Exponential tilting

It is elementary that if ρ is a Lévy density, corresponding to a density f for T, and b is a real number such that $\psi(b)$ defined by (17) is finite, then

$$\rho^{(b)}(x) = \rho(x)e^{-bx} \tag{41}$$

is also a Lévy density, and the corresponding density of T is

$$f^{(b)}(t) = f(t)\, e^{\psi(b)-bt} \tag{42}$$

It is also well known [34, Proposition 2.1.3] that if $\mathbb{P}^{(b)}$ denotes the probability distribution governing the Poisson set up with Lévy density $\rho^{(b)}$ then (42) extends to the absolute continuity relation

$$\frac{d\mathbb{P}^{(b)}}{d\mathbb{P}^{(0)}} = e^{\psi(b)-bT}. \tag{43}$$

This relation is equivalent to a combination of (42) and the following identity, which can also be verified using the construction of Lemma 2:

$$\text{PK}(\rho^{(b)} \,|\, t) = \text{PK}(\rho \,|\, t) \text{ for all } t > 0. \tag{44}$$

Consequently

$$\text{PK}(\rho^{(b)}, \gamma) = \text{PK}(\rho, \gamma) \tag{45}$$

for every γ. In particular, the distribution on $\mathcal{P}^{\downarrow}$ derived from the unconditioned Poisson model with Lévy density $\rho^{(b)}$ is

$$\mathrm{PK}(\rho^{(b)}) = \mathrm{PK}(\rho, \gamma^{(b)}) \tag{46}$$

where $\gamma^{(b)}$ is the $\mathbb{P}^{(b)}$ distribution of T, that is $\gamma^{(b)}(dt) = f^{(b)}(t)dt$ for $f^{(b)}$ as in (42). It can also be shown that if ρ' and ρ are two regular Lévy densities such that $\mathrm{PK}(\rho') = \mathrm{PK}(\rho, \gamma)$ for some γ, then $\rho' = \rho^{(b)}$ and $\gamma = \gamma^{(b)}$ for some b.

4.3 Deletion of Classes.

The following proposition, which generalizes a result of [41], provides motivation for study of $\mathrm{PK}(\rho, \gamma)$-partitions for other distributions γ besides $\gamma(dt) = f(t)dt$ corresponding to the unconditioned Poisson set up, and $\gamma = \delta_t$ corresponding to conditioning on $T = t$. Given a random partition Π of $\mathbb{N}$ with infinitely many classes, for each $k = 0, 1, \cdots$ let Π_k be the partition of $\mathbb{N}$ derived from Π by *deletion of the first k classes*, an operation made precise as follows. First let Π'_k be the restriction of Π to $H_k := \mathbb{N} - G_1 - \cdots - G_k$ where $G_1, \cdots G_k$ are the first k classes of Π in order of least elements, then derive Π_k on $\mathbb{N}$ from Π'_k on H_k by renumbering the points of H_k in increasing order.

Proposition 7 *Let Π be a $\mathrm{PK}(\rho, \gamma)$-partition of $\mathbb{N}$, and let Π_k be derived from Π be deletion of its first k classes. Then Π_k is a $\mathrm{PK}(\rho, \gamma_k)$-partition of $\mathbb{N}$, where $\gamma_k = \gamma Q^k$ for Q the Markov transition operator on $(0, \infty)$*

$$Q(t, dv) = \rho(t - v)(t - v)t^{-1} f(v) 1(0 < v < t) dv.$$

In particular, if Π is a $\mathrm{PK}(\rho)$ partition of $\mathbb{N}$, then Π_k is $\mathrm{PK}(\rho, \gamma_k)$-partition of $\mathbb{N}$, where γ_k is the distribution of T_k, the total sum of Poisson lengths T minus the sum of the first k lengths discovered by a process of length-biased sampling, as in (22).

Proof. According to a result of [41] which is implicit in Lemma 2, the sequence (T_k) is Markov with stationary transition probabilities given by Q. The conclusion follows from this observation, the construction of $\mathrm{PK}(\rho; \gamma)$, and the general construction of an exchangeable partition of $\mathbb{N}$ conditionally given its class frequencies [43].

5 Examples

5.1 The one-parameter Poisson-Dirichlet distribution.

Following Kingman [29], for the particular choice

$$\rho(x) = \theta\, x^{-1}\, e^{-bx} \tag{47}$$

where $\theta > 0$ and $b > 0$, corresponding to T with the gamma(θ, b) density

$$f(t) = \frac{b^\theta}{\Gamma(\theta)} t^{\theta-1} e^{-bt}, \tag{48}$$

the $\text{PK}(\rho)$ distribution is the *Poisson-Dirichlet distribution with parameter* θ, abbreviated $\text{PD}(\theta)$. Note the lack of dependence on the inverse scale parameter b. The well known fact the structural distribution of $\text{PD}(\theta)$ is beta$(1, \theta)$ follows immediately from (20). It follows easily from any one of the previous general formulas (27), (35), (36) or (40), that the EPPF of a $\text{PD}(\theta)$-partition $\Pi = (\Pi_n)$ is given by the formula

$$p_\theta(n_1, \cdots, n_k) = \frac{\theta^k \Gamma(\theta)}{\Gamma(\theta + n)} \prod_{i=1}^{k} (n_i - 1)! \qquad (n = \sum_{i=1}^{k} n_i). \tag{49}$$

This is a known equivalent [32, 43] of the Ewens sampling formula [18, 17] for the joint distribution of the number of blocks of Π_n of various sizes. It is also known [41, 49] that the following conditions on ρ are equivalent:

(i) ρ is of the form (47), for some $b > 0, \theta > 0$;

(ii) $\text{PK}(\rho \,|\, t) = \text{PK}(\rho)$ for all $t > 0$;

(iii) $\text{PK}(\rho) = \text{PD}(\theta)$ for some $\theta > 0$.

(iv) a $\text{PK}(\rho)$-partition has EPPF of the form (49) for some $\theta > 0$.

See also [4, 33] for further properties and applications of $\text{PD}(\theta)$.

5.2 Generalized gamma

After the one-parameter Poisson-Dirichlet family, the next simplest Lévy density ρ to consider is

$$\rho_{\alpha,c,b}(x) = c\, x^{-\alpha-1}\, e^{-bx} \tag{50}$$

for positive constants c and b, and α which is restricted to $0 \le \alpha < 1$ by the constraints on a Lévy density and (15). The corresponding distributions of T are known as *generalized gamma distributions* [8]. Note that the usual family of gamma distributions is recovered for $\alpha = 0$, and that a stable distribution with index α is obtained for $b = 0$ and $0 < \alpha < 1$. One can also take $\alpha = -\kappa$ for arbitrary $\kappa > 0$, except that in this model the Lévy measure has a total mass $\psi(\infty) < \infty$ so

$$\mathbb{P}(T = 0) = \exp(-\psi(\infty)) > 0,$$

contrary to the present assumption that the distribution of T has a density. Such models can be analyzed by first conditioning on the Poisson total number of lengths, which reduces the model to one with say m i.i.d. lengths with probability density proportional to ρ. In the case (50) for $\alpha = -\kappa$, that is to say that the lengths are i.i.d. gamma(κ, b) variables. This model for random partitions has been extensively studied. It is well known that features of the $\text{PD}(\theta)$ model can be derived by taking limits of this more elementary model with m i.i.d. gamma(κ, b) lengths as $\kappa \to 0$ and $m \to \infty$ with $\kappa m \to \theta$. See [45] for a review of this circle of ideas and its applications to species sampling models.

The $\mathrm{PK}(\rho_{\alpha,c,b})$ model for a random partition defined by $\rho_{\alpha,c,b}$ in (50) for $0 \le \alpha < 1$ was proposed by McCloskey [37], who first exploited the key idea of size-biased sampling in the setting of species sampling problems. Due to the remarks in Section 4 about scaling and exponential tilting, for $0 < \alpha < 1$ the family of $\mathrm{PK}(\rho_{\alpha,c,b}, \gamma)$ distributions, as γ varies over all distributions on $(0, \infty)$, depends only on α and not on c or b. So in studying this family of distributions on $\mathcal{P}^\downarrow$ and their associated exchangeable partitions of $\mathbb{N}$, the choice of c and b is entirely a matter of convenience. This study is taken up in the next section, with the choice of $b = 0$ and $c = \alpha/\Gamma(1-\alpha)$ which leads to the simplest form of most results. See also [8, 24, 23] regarding generalized gamma random measures and further developments.

5.3 The stable (α) model

Suppose now that $\mathbb{P}_\alpha$ governs the Poisson model for T with stable (α) distribution with Laplace transform

$$\mathbb{E}_\alpha[\exp(-\lambda T)] = \int_0^\infty e^{-\lambda x} f_\alpha(x)\, dx = \exp(-\lambda^\alpha) \tag{51}$$

for some $0 < \alpha < 1$, where $f_\alpha(x)$ is the stable(α) density of T, that is [52]

$$f_\alpha(t) = \frac{-1}{\pi} \sum_{k=0}^\infty \frac{(-1)^k}{k!} \sin(\pi\alpha k) \frac{\Gamma(\alpha k + 1)}{t^{\alpha k+1}}. \tag{52}$$

For $\alpha = \frac{1}{2}$ this reduces to the following formula of Doetsch [14, pp. 401-402] and Lévy [36]:

$$\mathbb{P}_{\frac{1}{2}}(2T \in dx)/dx = \tfrac{1}{2} f_{\frac{1}{2}}(\tfrac{1}{2}x) = \frac{1}{\sqrt{2\pi}} x^{-\frac{3}{2}} e^{-\frac{1}{2x}}. \tag{53}$$

Special results for $\alpha = \frac{1}{2}$, discussed in Section 8, involve cancellations due to simplification of $f_\alpha(pt)/f_\alpha(t)$ for $0 < p < 1$, which does not appear to be possible for general α. The Lévy density corresponding to the Laplace transform (51) is well known to be

$$\rho_\alpha(x) = \frac{\alpha\, x^{-\alpha-1}}{\Gamma(1-\alpha)} \qquad (x > 0). \tag{54}$$

Write $\mathbb{P}_\alpha(\,\cdot\,|\, t)$ for $\mathbb{P}_\alpha(\,\cdot\,|\, T = t)$. So the $\mathbb{P}_\alpha$ distribution of (P_i) on $\mathcal{P}^\downarrow$ is $\mathrm{PK}(\rho_\alpha)$, and the $\mathbb{P}_\alpha(\,\cdot\,|\, t)$ distribution of (P_i) is $\mathrm{PK}(\rho_\alpha \,|\, t)$. Note from (51) that if T_c is the total length in the model governed by $c\rho_\alpha$ for a constant $c > 0$, then T_c has the same distribution as $c^{1/\alpha} T_1$ for $T_1 = T$ as in (51). Together with similar scaling properties of the lengths J_i, this implies that for all $0 < \alpha < 1$ and $t > 0$ there is the formula

$$\mathrm{PK}(c\rho_\alpha \,|\, t) = \mathrm{PK}(\rho_\alpha \,|\, c^{-1/\alpha}\, t). \tag{55}$$

Formulas for the $\text{PK}(\rho_\alpha \,|\, t)$ distribution are described in Section 5.4. These formulas can be understood as disintegrations of simpler formulas obtained in [43], and recalled in Section 6, for a particular subfamily of the class of $\text{PK}(\rho_\alpha, \gamma)$ distributions.

One reason for special interest in the Kingman family associated with the stable Lévy densities ρ_α is the following result which will be proved elsewhere.

Theorem 8 *The EPPF of an exchangeable random partition Π of $\mathbb{N}$ with an infinite number of classes with proper frequencies has an EPPF of the Gibbs form*

$$p(n_1, \cdots, n_k) = c_{n,k} \prod_{i=1}^{k} w_{n_i} \text{ where } n = \textstyle\sum_{i=1}^{k} n_i \tag{56}$$

for some positive weights $w_1 = 1, w_2, w_3, \ldots$ and some $c_{n,k}$ if and only if

$$w_m = \prod_{j=1}^{m-1} (j - \alpha) \qquad (m = 1, 2, \ldots)$$

for some $0 \le \alpha < 1$. If $\alpha = 0$ then the distribution of Π corresponds to $\int_0^\infty \text{PD}(\theta)\gamma(d\theta)$ for some probability distribution γ on $(0, \infty)$, whereas if $0 < \alpha < 1$ then the distribution of Π corresponds to $\text{PK}(\rho_\alpha, \gamma) := \int_0^\infty \text{PK}(\rho_\alpha \,|\, t)\gamma(dt)$ for some γ.

See also Kerov [28] and Zabell [57] for related characterizations of the two-parameter family discussed in Section 6. This family is characterized by an EPPF of the form (56) with $c_{n,k}$ a product of a function of n and a function of k.

5.4 Conditioning on T

Assume throughout this section that $0 < \alpha < 1$. Immediately from (20) and (54), in the $\text{PK}(\rho_\alpha \,|\, t)$ model, the distribution of $\tilde{P}_1$ has density

$$\tilde{f}_\alpha(p \,|\, t) = \frac{\alpha (pt)^{-\alpha}}{\Gamma(1-\alpha)} \frac{f_\alpha((1-p)t)}{f_\alpha(t)} \qquad (0 < p < 1). \tag{57}$$

Let h be a non-negative measurable function with $\mathbb{E}_\alpha h(T) = \int_0^\infty h(t) f_\alpha(t) dt = 1$, and let $h \cdot f_\alpha$ denote the distribution on $(0, \infty)$ with density $h(t) f_\alpha(t)$. Then by integration from (57), under the probability $\mathbb{P}_{\alpha,h}$ governing the $\text{PK}(\rho_\alpha, h \cdot f_\alpha)$ model, the structural distribution of $\tilde{P}_1$ has density

$$\mathbb{P}_{\alpha,h}(\tilde{P}_1 \in dp)/dp = \frac{\alpha}{\Gamma(1-\alpha)} p^{-\alpha} (1-p)^{\alpha-1} \eta_{\alpha,h}(1-p) \quad (0 < p < 1) \tag{58}$$

where

$$\eta_{\alpha,h}(u) := \int_0^\infty v^{-\alpha} h(v/u) f_\alpha(v) dv = \mathbb{E}_\alpha[T^{-\alpha} h(T/u)]. \tag{59}$$

For instance, it is known [41] that

$$C_{\alpha,\theta} := \mathbb{E}_\alpha(T^{-\theta}) = \frac{\Gamma(\frac{\theta}{\alpha}+1)}{\Gamma(\theta+1)} \qquad (\theta > -\alpha). \tag{60}$$

So for $\theta > -\alpha$, (58) and (59) imply:

$$\text{if } h(t) = C_{\alpha,\theta}^{-1} t^{-\theta} \text{ then } \tilde{P}_1 \text{ has beta}(1-\alpha, \alpha+\theta) \text{ distribution.} \tag{61}$$

This example is discussed further in the next section. As another example, if $h(t) = \exp(b^\alpha - bt)$ for some $b > 0$, then according to (46) the model $\text{PK}(\rho_\alpha, h \cdot f_\alpha)$ is identical to the unconditioned generalized gamma model $\text{PK}(\rho_{\alpha,b})$ with

$$\rho_{\alpha,b}(x) := \rho_\alpha(x) e^{-bx} = \frac{\alpha}{\Gamma(1-\alpha)} \frac{e^{-bx}}{x^{\alpha+1}} \qquad (x > 0).$$

So the structural density of the $\text{PK}(\rho_{\alpha,b})$ model is given by formula (58) with

$$\eta_{\alpha,h}(u) = \exp(b^\alpha) \mathbb{E}_\alpha[T^{-\alpha} \exp(-bT/u)]. \tag{62}$$

For $\alpha = \frac{1}{2}$ the expectation in (62) can be evaluated by using (53) to write for $\xi > 0$

$$\mathbb{E}_{\frac{1}{2}}[T^{-\frac{1}{2}} \exp(-\xi T)] = \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{dx}{x^2}\, e^{-(\xi x + 1/x)/2} = 2\sqrt{\frac{\xi}{\pi}} K_1(\sqrt{\xi}) \tag{63}$$

where K_1 is the usual modified Bessel function. Thus for $b > 0$ the $\text{PK}(\rho_{\frac{1}{2},b})$ model associated with the inverse Gaussian distribution [54] has structural distribution with density $\tilde{f}_{\frac{1}{2},b}$ given by the formula

$$\tilde{f}_{\frac{1}{2},b}(p) = \frac{\sqrt{b} e^{\sqrt{b}}}{\pi \sqrt{p}(1-p)} K_1 \left(\sqrt{\frac{b}{(1-p)}} \right) \qquad (0 < p < 1). \tag{64}$$

Proposition 9 *For $0 < \alpha < 1, q > 0$ let $\mu_\alpha(q \,|\, t)$ denote the qth moment of the structural density* (57) *of the $\text{PK}(\rho_\alpha \,|\, t)$ distribution:*

$$\mu_\alpha(q \,|\, t) := \int_0^1 p^q \tilde{f}_\alpha(p \,|\, t)\, dp = \mathbb{E}_\alpha(\tilde{P}_1^q \,|\, t). \tag{65}$$

Then for each $t > 0$ the EPPF of a $\text{PK}(\rho_\alpha \,|\, t)$ partition of $\mathbb{N}$ is

$$p_\alpha(n_1, \cdots, n_k \,|\, t) = \frac{\Gamma(1-\alpha)}{\Gamma(n-k\alpha)} \left(\frac{\alpha}{t^\alpha}\right)^{k-1} \mu_\alpha(n-1-k\alpha+\alpha \,|\, t) \prod_{i=1}^k [1-\alpha]_{n_i-1} \tag{66}$$

where

$$[1-\alpha]_{n_i-1} := \prod_{j=1}^{n_i-1} (j-\alpha) = \frac{\Gamma(n_i-\alpha)}{\Gamma(1-\alpha)}.$$

Alternatively,

$$p_\alpha(n_1,\cdots,n_k\,|\,t) = \frac{\alpha^k}{t^n} g_\alpha(n-k\alpha\,|\,t)\prod_{i=1}^{k}[1-\alpha]_{n_i-1} \tag{67}$$

where $g_\alpha(q\,|\,t) := (\Gamma(q)f_\alpha(t))^{-1}\int_0^t f_\alpha(t-v)v^{q-1}dv$.

Proof. This is read from Theorem 4, since the integral (28) reduces to a standard Dirichlet integral. □

As checks on (66), the symmetry in $(n_1,\ldots,n_k)$ is still evident, and $p_\alpha(n\,|\,t) = \mu_\alpha(n-1\,|\,t)$ as required by (8). However, the addition rules (5) for this EPPF are not at all obvious. Rather, they amount to the following identity involving moments of the structural distribution:

Corollary 10 *The moments $\mu_\alpha(q|t)$ of the structural distribution on $(0,1)$ associated with the* PK$(\rho_\alpha\,|\,t)$ *distribution on $\mathcal{P}^\downarrow$ satisfy the following identity: for all $1\le k\le n$ and $t>0$*

$$\mu_\alpha(n-1-k\alpha+\alpha\,|\,t) = \mu_\alpha(n-k\alpha+\alpha\,|\,t) + \frac{\Gamma(n-k\alpha)\,\alpha\,t^{-\alpha}}{\Gamma(n+1-k\alpha-\alpha)}\,\mu_\alpha(n-k\alpha\,|\,t). \tag{68}$$

To illustrate, according to the simplest addition rule (6),

$$1 = p_\alpha(2\,|\,t) + p_\alpha(1,1\,|\,t),$$

which amounts to (68) for $n=k=1$, that is

$$1 = \mu_\alpha(1\,|\,t) + \frac{\Gamma(1-\alpha)}{\Gamma(2-2\alpha)}\,\frac{\alpha}{t^\alpha}\,\mu_\alpha(1-\alpha\,|\,t). \tag{69}$$

The addition rule underlying (68) can be checked for general α by an argument described in Section 6. In the case $\alpha=\frac{1}{2}$, the later formulae (99) and (93) show that (68) reduces to a known recursion (106) for the Hermite function.

Repeated application of (68) shows that for each $1\le k\le n$ the moment on the left side of (66) can be expressed as a linear combination of integer moments $\mu_\alpha(j\,|\,t)$ for $j=0,\cdots,n-1$, with coefficients depending on n,k,α,t which could easily be computed recursively. But except in the special case $\alpha=\frac{1}{2}$ discussed in Section 8, even the integer moments seem difficult to evaluate.

6 The two-parameter Poisson-Dirichlet family

For $0<\alpha<1, \theta>-\alpha$, let $\gamma_{\alpha,\theta}$ denote the distribution on $(0,\infty)$ with density $C_{\alpha,\theta}^{-1}t^{-\theta}$ at t relative to the stable(α) distribution of T defined by (51), that is

$$\gamma_{\alpha,\theta}(dt) = C_{\alpha,\theta}^{-1}\,t^{-\theta}\,f_\alpha(t)\,dt \tag{70}$$

where $C_{\alpha,\theta} := \mathbb{E}_\alpha(T^{-\theta}) = \Gamma(\frac{\theta}{\alpha}+1)/\Gamma(\theta+1)$ as in (60) and (61).

Definition 11 [41, 50] The *Poisson-Dirichlet distribution with two parameters* (α, θ), denoted $\text{PD}(\alpha, \theta)$, is the distribution on $\mathcal{P}^{\downarrow}$ defined for $0 \le \alpha < 1, \theta > -\alpha$ by

$$\text{PD}(\alpha,\theta) = \begin{cases} \text{PD}(\theta) & \text{for } \alpha = 0, \theta > 0 \\ \text{PK}(\rho_\alpha, \gamma_{\alpha,\theta}) & \text{for } 0 < \alpha < 1, \theta > -\alpha \end{cases} \tag{71}$$

This family of distributions on $\mathcal{P}^{\downarrow}$ has some remarkable properties and applications. As shown in [41], it follows from Lemma 2 that if (P_i) has $\text{PD}(\alpha,\theta)$distribution then the corresponding size-biased sequence $(\tilde{P}_j)$ can be represented as

$$\tilde{P}_j = W_j \prod_{i=1}^{j-1} (1 - W_i) \tag{72}$$

where the W_j are independent with beta$(1-\alpha, \theta + j\alpha)$ distributions. (73)

So the $\text{PD}(\alpha,\theta)$ distribution can just as well be defined, without reference to the Poisson-Kingman construction, as the distribution of (P_i) defined by ranking $(\tilde{P}_j)$ constructed by (72) from independent W_j as in (72). The sequence $(\tilde{P}_j)$ defined by (72) and (73) for $0 \le \alpha < 1$ and $\theta > 0$ was considered by Engen [15] as a model for species abundances. See [50] for further study of the $\text{PD}(\alpha,\theta)$ family. It was shown in [44] that if (P_i) is a random element of $\mathcal{P}^{\downarrow}$ with $P_i > 0$ a.s. for all i and the corresponding size-biased sequence $(\tilde{P}_j)$ admits the representation (72) with independent residual fractions W_j, then the W_j must have beta distributions as described in (73), and hence the distribution of (P_i) must be $\text{PD}(\alpha,\theta)$ for some $0 \le \alpha < 1$ and $\theta > -\alpha$. Reformulated in terms of random partitions, and combined with Proposition 7, this yields the following:

Proposition 12 *Let* Π *be the exchangable random partition of* $\mathbb{N}$ *derived by sampling from a random element* (P_i) *of* $\mathcal{P}^{\downarrow}$ *with* $P_i > 0$ *for all* i*. Let* Π_k *be derived from* Π *by deletion of the first* k *classes of* Π*, with classes in order of appearance, as defined above Proposition* 7*. Then the following are equivalent*
(i) *for each* k*,* Π_k *is independent of the frequencies* $(\tilde{P}_1, \cdots, \tilde{P}_k)$ *of the first* k *classes of* Π*;*
(ii) Π *is a* $\text{PD}(\alpha,\theta)$*-partition for some* $0 \le \alpha < 1$ *and* $\theta > -\alpha$*, in which case* Π_k *is a* $\text{PD}(\alpha, \theta + k\alpha)$*-partition.*

As shown in [43], the independence property (72) of the residual fractions W_j of a $\text{PD}(\alpha,\theta)$-partition allows the corresponding EPPF $p_{\alpha,\theta}(n_1, \ldots, n_k)$ to be evaluated using (3). The result is as follows. For all $0 \le \alpha < 1$ and $\theta > -\alpha$,

$$p_{\alpha,\theta}(n_1, \ldots, n_k) = \frac{[\theta+\alpha]_{k-1;\alpha}}{[\theta+1]_{n-1}} \prod_{i=1}^{k} [1-\alpha]_{n_i - 1} \tag{74}$$

where $n = \sum_{i=1}^{k} n_i$ and for real x and a and non-negative integer m

$$[x]_{m;a} = \begin{cases} 1 \ \text{ for } \ m = 0 \\ x(x+a)\cdots(x+(m-1)a) \ \text{ for } \ m = 1, 2, \ldots \end{cases}$$

and $[x]_m = [x]_{m;1}$. The previous formula (49) is the special case of (74) for $\alpha = 0$. Both this case of (74), and the case when $0 < \alpha < 1$ and $\theta = 0$, follow easily from (36). Formula (74) shows that a $\text{PD}(\alpha, \theta)$ partition Π of $\mathbb{N}$ to be constructed sequentially as follows [43, 45]. Starting from $\Pi_1 = \{\{1\}\}$, given that Π_n has been constructed as a partition of $\mathbb{N}_n$ with say k blocks of sizes $(n_1, \cdots, n_k)$, define Π_{n+1} by assigning the new element $n + 1$ to the jth class whose current size is n_j with probability

$$\mathbb{P}(j \uparrow \,|\, n_1, \cdots, n_k) = \frac{n_j - \alpha}{n + \theta} \tag{75}$$

for $1 \le j \le k$, and assigning $n+1$ to a new class numbered $k+1$ with the remaining probability

$$\mathbb{P}(k + 1 \uparrow \,|\, n_1, \cdots, n_k) = \frac{k\alpha}{n + \theta} \tag{76}$$

For $\alpha = 0$ and $\theta > 0$ this is generalization of Polya's urn scheme developed by Blackwell-McQueen [7] and Hoppe [21]. See [43, 45, 20] for consideration of more general prediction rules for exchangeable random partitions.

The following calculation shows how to derive either of the two EPPF's (74) and (66) from the other. The argument also shows that the function $p_\alpha(n_1, \cdots, n_k \,|\, t)$ defined by (66) satisfies the addition rules of an EPPF as a consequence of the corresponding addition rules for $p_{\alpha,\theta}(n_1, \ldots, n_k)$, which are much more obvious.

The kernel $\gamma_{\alpha,\theta}(dt)$ introduced in (70), is now viewed for a fixed α as a family of probability distributions on $(0, \infty)$ indexed by $\theta \in (-\alpha, \infty)$, that is a Markov kernel γ_α from $(-\alpha, \infty)$ to $(0, \infty)$. For a non-negative measurable function $h = h(t)$ with domain $(0, \infty)$, define a function $\gamma_\alpha h = (\gamma_\alpha h)(\theta)$ with domain $(-\alpha, \infty)$ by the usual action of this Markov kernel as an integral operator:

$$(\gamma_\alpha h)(\theta) = \int_0^\infty \gamma_{\alpha,\theta}(dt) h(t) \tag{77}$$

Then say $(\gamma_\alpha h)(\theta)$ is the *γ_α-transform* of $h(t)$. Let $\mathbb{E}_{\alpha,\theta}$ denote expectation with respect to the probability distribution

$$\mathbb{P}_{\alpha,\theta}(\cdot) := \int_0^\infty \mathbb{P}_\alpha(\cdot \,|\, t) \gamma_{\alpha,\theta}(dt).$$

By definition, for each non-negative random variable X governed by the family of conditional laws $(\mathbb{P}_\alpha(\cdot \,|\, t), t > 0)$,

$$\text{the } \gamma_\alpha\text{-transform of } \mathbb{E}_\alpha(X \,|\, t) \text{ is } \mathbb{E}_{\alpha,\theta}(X). \tag{78}$$

In particular, for each $(n_1, \ldots, n_k)$,

$$\text{the } \gamma_\alpha\text{-transform of } p_\alpha(n_1, \cdots, n_k \,|\, t) \text{ is } p_{\alpha,\theta}(n_1, \ldots, n_k). \tag{79}$$

An obvious change of variable allows uniqueness and inversion results for the γ_α-transform to be deduced from standard results for Mellin or bilateral exponential transforms. So the problem is just to show that the γ_α-transform of the right side of (66) is the right side of (74). To see this, observe first that for each $q > 0$, because $\mu_\alpha(q|t) := \mathbb{E}_\alpha(\tilde{P}_1^q \,|\, t)$,

$$\text{the } \gamma_\alpha\text{-transform of } \mu_\alpha(q|t) \text{ is } \mathbb{E}_{\alpha,\theta}(\tilde{P}_1^q) = \frac{\Gamma(1-\alpha+q)\Gamma(1+\theta)}{\Gamma(1+\theta+q)\Gamma(1-\alpha)} \tag{80}$$

where $\mathbb{E}_{\alpha,\theta}(\tilde{P}_1^q)$ is evaluated using (61). To deal with the factor of $t^{-(k-1)\alpha}$ in (66), note from (60) that for each $\beta > 0$, and any $h(t)$,

$$\text{the } \gamma_\alpha\text{-transform of } t^{-\beta}h(t) \text{ is } \frac{\Gamma(\frac{\theta}{\alpha} + \frac{\beta}{\alpha} + 1)\Gamma(\theta+1)}{\Gamma(\frac{\theta}{\alpha}+1)\Gamma(\theta+\beta+1)}\,(\gamma_\alpha h)(\theta+\beta). \tag{81}$$

By (80) for $q = n - 1 - k\alpha + \alpha$ and (81) for $\beta = \alpha k - \alpha$ and $h(t) = \mu_\alpha(q \,|\, t)$ the right side of (66) has for its γ_α-transform the following function of θ:

$$\frac{\alpha^{k-1}\Gamma(1-\alpha)}{\Gamma(n-k\alpha)}\,\frac{\Gamma(\frac{\theta}{\alpha}+k)\Gamma(\theta+1)}{\Gamma(\frac{\theta}{\alpha}+1)\Gamma(\theta+k\alpha-\alpha+1)}\,\frac{\Gamma(n-k\alpha)\Gamma(1+\theta+k\alpha-\alpha)}{\Gamma(n+\theta)\Gamma(1-\alpha)}\prod_{i=1}^{k}[1-\alpha]_{n_i-1}$$

which reduces by cancellation to the right side of (74).

6.1 The α-diversity

Let Π be an exchangeable random partition of $\mathbb{N}$ with ranked frequencies (P_i). Let K_n denote the number of classes of Π_n, the partition of $\mathbb{N}_n$ induced by Π. Say that Π has *α-diversity* S and write α-DIVERSITY$(\Pi) = S$ iff there exists a random variable S with $0 < S < \infty$ a.s. and

$$K_n \sim S n^\alpha \text{ as } n \to \infty \tag{82}$$

where for two sequences of random variables A_n and B_n, the notation $A_n \sim B_n$ will now be used to indicate that $A_n/B_n \to 1$ almost surely as $n \to \infty$. According to a result of Karlin [27], applied conditionally given (P_i), if these ranked frequencies are such that

$$P_i \sim \left(\frac{S}{\Gamma(1-\alpha)i}\right)^{\frac{1}{\alpha}} \tag{83}$$

for some $0 < S < \infty$ then Π has α-diversity S.

Proposition 13 *Suppose* Π *is a* PK(ρ_α, γ) *partition of* $\mathbb{N}$ *for some* $0 < \alpha < 1$ *and some probability distribution* γ *on* $(0, \infty)$*. Then*
(i) α-DIVERSITY$(\Pi) = S$ for a random variable S with $S = T^{-\alpha}$ where $T = S^{-1/\alpha}$ has distribution γ. In particular, $S = t^{-\alpha}$ is constant if Π is a PK$(\rho_\alpha \,|\, t)$ partition.
(ii) A regular conditional distribution for Π given $S = s$ is defined by the EPPF $p_\alpha(n_1, \cdots, n_k | s^{-1/\alpha})$ obtained by setting $t = s^{-1/\alpha}$ in (66).
(iii) In particular, both (i) and (ii) hold if Π is a PD(α, θ) partition for some $\theta > -\alpha$. Then the α-diversity S of Π is $S = T^{-\alpha}$ for T with the distribution $\gamma_{\alpha,\theta}$ defined by (70).

Proof. Suppose that (P_i) has PK(ρ_α, γ) distribution. The fact that (83) holds for $S = T^{-\alpha}$ in the unconditioned case where T has stable(α) distribution is due to Kingman [29]. Kingman's argument, using the law of large numbers for small jumps of the Poisson process, applies just as well for T conditioned to be a constant t. So (83) follows in general by mixing over t. □

See [50] and papers cited there for further information about the Mittag-Leffler distribution of $S = T^{-\alpha}$ derived from a PD$(\alpha, 0)$ partition. The corresponding distribution of S for PD(α, θ) has density at s proportional to $s^{\theta/\alpha}$ relative to this Mittag-Leffler distribution.

As shown in [50, Proposition 10], if Π is a partition of $\mathbb{N}$ whose ranked frequencies (P_i) have the PD$(\alpha, 0)$ distribution, then $S = \alpha$-DIVERSITY(Π) can be recovered from Π or (P_i) via either (81) or (83). Then $T = S^{-1/\alpha}$ has stable(α) distribution as in (51), and (TP_i) is then sequence of points of a Poisson process with Lévy density ρ_α. See also [47, 48] for more about the distribution of K_n derived from a PD(α, θ) partition.

7 Application to lengths of excursions.

This section reviews some results of [41, 49, 46, 50]. Let $\mathbb{P}^0_\alpha$ govern a strong Markov process B starting at a recurrent point 0 of its statespace, such that the inverse $(\tau_\ell, \ell \geq 0)$ of the local time process $(L_t, t \geq 0)$ of B at zero is a stable subordinator of index α for some $0 < \alpha < 1$. That is to say, $\mathbb{E}^0_\alpha \exp(-\lambda \tau_1) = \exp(-c\lambda^\alpha)$ for some constant $c > 0$. So the $\mathbb{P}^0_\alpha$ distribution of τ_1 is the $\mathbb{P}_\alpha$ distribution of $c^{1/\alpha} T$ for T as in (51). For example, B could be a one-dimensional Brownian motion ($\alpha = \frac{1}{2}$) or Bessel process of dimension $2 - 2\alpha$. In the Brownian case, take $c = \sqrt{2}$ to obtain the usual normalization of local time as occupation density relative to Lebesgue measure, which makes $L_1 \stackrel{d}{=} |B_1|$. Let $M = \{t : 0 \leq t \leq 1, B_t = 0\}$ denote the random closed subset of $[0, 1]$ defined by the zero set of B. Component intervals of the complement of M relative to $[0, 1]$ are called *excursion intervals.* For $0 \leq t \leq 1$ let $G_t = \sup\{M \cap [0, t]\}$, the last zero of B before time t. Note that with probability one, $G_1 < 1$, so one of the excursion intervals is the *meander interval* $(G_1, 1]$, whose length $1 - G_1$ is one of the lengths appearing in the list (P_i) say of ranked lengths of excursion intervals. According to the main result of [49],

$$\text{the sequence } (P_i) \text{ of ranked lengths has } \mathrm{PD}(\alpha, 0) \text{ distribution} \tag{84}$$

Let $U_1, U_2, \cdots$ be a sequence of i.i.d. uniform $[0,1]$ random variables, independent of B, called the sequence of *sample points.* Let $\Pi = (\Pi_n)$ be the random partition of $\mathbb{N}$ generated by the random equivalence relation $i \sim j$ iff $G_{U_i} = G_{U_j}$. That is to say $i \sim j$ iff U_i and U_j fall in the same excursion interval. So for example $\Pi_5 = \{\{1,2,5\},\{3\},\{4\}\}$ iff U_1, U_2 and U_5 fall in one excursion interval, U_3 in another, and U_4 in a third. By translation of results of [49, 50] into present notation

$$\Pi \text{ is a } \mathrm{PD}(\alpha, 0) \text{ partition and } \alpha\text{-DIVERSITY}(\Pi) = cL_1 \tag{85}$$

where L_1 is the local time of B at zero up to time 1. By construction, the sequence $(\tilde{P}_j)$ of class frequencies of Π is the sequence of lengths of excursion intervals in the order they are discovered by the sample points, and (P_i) is recovered from $(\tilde{P}_j)$ by ranking. To illustrate formula (74), U_1 and U_2 fall in different excursion intervals with probability $p_{\alpha,0}(1,1) = \alpha$, and in the same one with probability $p_{\alpha,0}(2) = 1 - \alpha$. Similarly, given that the local time is $L_1 = \ell$, two sample points fall in the same excursion interval with probability $p_\alpha(2 \,|\, (c\ell)^{-1/\alpha})$, and in different excursion intervals with probability $p_\alpha(1,1 \,|\, (c\ell)^{-1/\alpha})$, for $p_\alpha(\cdots \,|\, t)$ defined by (66). See Section 8 for evaluation of these functions in the case $\alpha = \frac{1}{2}$ corresponding to a Brownian motion B.

Let $R_n = 1 - \tilde{P}_1 - \cdots - \tilde{P}_n$, which is the total length of excursions which remain undiscovered after the sampling process has found n distinct excursion intervals. The result of Proposition 12 in this setting, due to [41], is that for each $n = 0, 1, 2, \cdots$ a $\mathrm{PD}(\alpha, n\alpha)$ distributed sequence is obtained by ranking the sequence

$$\frac{1}{R_n}(\tilde{P}_{n+1}, \tilde{P}_{n+2}, \cdots) \tag{86}$$

of relative excursion lengths which remain after discovery of the first n intervals. For $n = 1$ the same $\mathrm{PD}(\alpha, \alpha)$ distribution is obtained more simply by deleting the meander of length $1 - G_1$, renormalizing and reranking. This is due to the result of [49] that the length $1 - G_1$ of the meander interval is a size-biased choice from (P_i). As the excursion lengths in this case are just the excursion lengths of a standard bridge, equivalent to conditioning on $B_1 = 0$, the ranked excursion lengths of such a bridge have $\mathrm{PD}(\alpha, \alpha)$ distribution. As first shown in [49], this implies that both the unconditioned process B and the bridge B given $B_1 = 0$ share a common conditional distribution for the ranked excursion lengths (P_i) given the local time L_1. In present notation, this conditional distribution of (P_i) given $L_1 = \ell$, with or without conditioning on $B_1 = 0$, is $\mathrm{PK}(\rho_\alpha | (c\ell)^{-1/\alpha})$.

One final identity is worth noting. As a consequence of the above discussion, for the process B, the conditional distribution of the meander length $1-G_1$ given $L_1=\ell$ is given by

$$\mathbb{P}^0_\alpha(1-G_1\in dp|L_1=\ell)=\mathbb{P}^0_\alpha(\tilde{P}_1\in\ dp|L_1=\ell)=\tilde{f}_\alpha(p|(c\ell)^{-1/\alpha})dp \tag{87}$$

where $\tilde{f}_\alpha(p|t)$ as in (57) is the structural density of the Poisson model for stable (α) distributed T conditioned on $T=t$. So the moment function $\mu_\alpha(q\,|\,t)$ appearing in the EPPF (66) of this model can be interpreted in the present setting as

$$\mu_\alpha(q\,|\,t)=\mathbb{E}^0_\alpha[(1-G_1)^q\,|\,L_1=c^{-1}t^{-\alpha}]. \tag{88}$$

8 The Brownian excursion partition

In this section let Π be the *Brownian excursion partition*, that is the random partition of $\mathbb{N}$ generated by uniform random sampling of points from the interval $[0,1]$ partitioned by the excursion intervals of a standard Brownian motion B. According to the result of [49] recalled in (84),

$$\Pi \text{ is a } \mathrm{PK}(\rho_{\frac{1}{2}})=\mathrm{PD}(\tfrac{1}{2},0) \text{ partition.} \tag{89}$$

With conditioning on $B_1=0$, the process B becomes a standard Brownian bridge. So Π given $B_1=0$ is a $\mathrm{PD}(\frac{1}{2},\frac{1}{2})$ partition, as discussed in the previous subsection. Features of the distribution of Π and the conditional distribution of Π given $B_1=0$ were described in [46]. This section presents refinements of these results obtained by conditioning on L_1, the local time of B at 0 up to time 1, with the usual normalization of Brownian local time as occupation density relative to Lebesgue measure. Unconditionally, L_1 has the same distribution as $|B_1|$, that is

$$\mathbb{P}(L_1\in d\lambda)=\mathbb{P}(|B_1|\in d\lambda)=2\varphi(\lambda)d\lambda \qquad\qquad (\lambda>0)$$

where $\varphi(z):=(1/\sqrt{2\pi})\exp(-\frac{1}{2}z^2)$ is the standard Gaussian density of B_1. Whereas the conditional distribution of L_1 given $B_1=0$ is the Rayleigh distribution

$$\mathbb{P}(L_1\in d\lambda\,|\,B_1=0)=\sqrt{2\pi}\lambda\varphi(\lambda)d\lambda \qquad\qquad (\lambda>0).$$

Note from (85) that the $\frac{1}{2}$-diversity of Π is the random variable $\sqrt{2}L_1$. So the number K_n of blocks of Π grows almost surely like $\sqrt{2n}L_1$ as $n\to\infty$. For $\lambda\geq 0$ let $\Pi(\lambda)$ denote a random partition with

$$\Pi(\lambda)\overset{d}{=}(\Pi\,|\,L_1=\lambda)\overset{d}{=}(\Pi\,|\,L_1=\lambda,B_1=0) \tag{90}$$

where $\overset{d}{=}$ denotes equality in distribution. So according to the previous discussion,

$$\Pi(\lambda) \text{ is a } \mathrm{PK}(\rho_{\frac{1}{2}}\,|\,\tfrac{1}{2}\lambda^{-2}) \text{ partition} \tag{91}$$

whose $\frac{1}{2}$-diversity is $\sqrt{2}\lambda$. Let $\text{PD}(\frac{1}{2} \,||\, \lambda)$ denote the probability distribution on $\mathcal{P}^{\downarrow}$ associated with $\Pi(\lambda)$, that is the common distribution of ranked lengths of excursions of a Brownian motion or Brownian bridge over $[0,1]$ given $L_1 = \lambda$. Then by Definition 11 and (53), for $\theta > -\frac{1}{2}$ there is the identity of probability laws on $\mathcal{P}^{\downarrow}$

$$\text{PD}(\tfrac{1}{2}, \theta) = \frac{2}{\mathbb{E}[|B_1|^{2\theta}]} \int_0^\infty \text{PD}(\tfrac{1}{2} \,||\, \lambda) \lambda^{2\theta} \varphi(\lambda) d\lambda \tag{92}$$

where, according to the gamma($\frac{1}{2}$) distribution of $\frac{1}{2}B_1^2$ and the duplication formula for the gamma function,

$$\mathbb{E}(|B_1|^{2\theta}) = 2^\theta \frac{\Gamma(\theta + \frac{1}{2})}{\Gamma(\frac{1}{2})} = 2^{-\theta} \frac{\Gamma(2\theta+1)}{\Gamma(\theta+1)} \qquad (\theta > -\tfrac{1}{2}). \tag{93}$$

It was shown in [3] (see also [5, 48]) that it is possible to construct the Brownian excursion partitions as a partition valued *fragmentation process* $(\Pi(\lambda), \lambda \geq 0)$, meaning that $\Pi(\lambda)$ is constructed for each λ on the same probability space, in such a way that $\Pi(\lambda)$ is a coarser partition than $\Pi(\mu)$ whenever $\lambda < \mu$. The question of whether a similar construction is possible for index α instead of index $\frac{1}{2}$ remains open. A natural guess is that such a construction might be made with one of the self-similar fragmentation processes of Bertoin [6], but Miermont and Schweinsberg [38] have recently shown that a construction of this form is possible only for $\alpha = \frac{1}{2}$.

8.1 Length biased sampling

Let $\tilde{P}_j(\lambda)$ denote the frequency of the jth class of $\Pi(\lambda)$. So $(\tilde{P}_j(\lambda), j = 1, 2 \ldots)$ is distributed like the lengths of excursions of B over $[0,1]$ given $L_1 = \lambda$, as discovered by a process of length-biased sampling. In view of Lévy's formula (53) for the stable($\frac{1}{2}$) density, the formula (57) reduces for $\alpha = \frac{1}{2}$ to the following more explicit formula for the structural density of $\Pi(\lambda)$:

$$\mathbb{P}(\tilde{P}_1(\lambda) \in dp) = \frac{\lambda}{\sqrt{2\pi}} p^{-\frac{1}{2}} (1-p)^{-\frac{3}{2}} \exp\left(-\frac{\lambda^2}{2} \frac{p}{(1-p)} \right) dp \qquad (0 < p < 1) \tag{94}$$

or equivalently

$$\mathbb{P}(\tilde{P}_1 \leq y) = 2\Phi\left(\lambda \sqrt{\frac{y}{1-y}} \right) - 1 \qquad (0 \leq y < 1) \tag{95}$$

where $\Phi(z) := \mathbb{P}(B_1 \leq z)$ is the standard Gaussian distribution function. Put another way, there is the equality in distribution

$$\tilde{P}_1(\lambda) \stackrel{d}{=} \frac{B_1^2}{\lambda^2 + B_1^2}. \tag{96}$$

Furthermore, by a similar analysis using Lemma 1, there is the following result which shows how to construct the whole sequence $(\tilde{P}_j(\lambda), j \geq 1)$ for any $\lambda > 0$ from a single sequence of independent standard Gaussian variables. Then $\Pi(\lambda)$ can be constructed by sampling from $(\tilde{P}_j(\lambda), j \geq 1)$ as discussed in Section 2.

Proposition 14 [3, Corollary 5] *Fix $\lambda > 0$. A sequence $(\tilde{P}_j(\lambda), j \geq 1)$ is distributed like a length-biased random permutation of the lengths of excursions of a Brownian motion or standard Brownian bridge over $[0,1]$ conditioned on $L_1 = \lambda$ if and only if*

$$\tilde{P}_j(\lambda) = \frac{\lambda^2}{\lambda^2 + S_{j-1}} - \frac{\lambda^2}{\lambda^2 + S_j} \tag{97}$$

where $S_j := \sum_{i=1}^{j} X_i$ for X_i which are independent and identically distributed like B_1^2 for a standard Gaussian variable B_1.

Let $\mu(q \,||\, \lambda)$ denote the qth moment of the distribution of $\tilde{P}_1(\lambda)$. So in the notation of (65) and (68)

$$\mu(q \,||\, \lambda) := \mathbb{E}[(\tilde{P}_1(\lambda))^q] = \mu_{\frac{1}{2}}(q \,|\, \tfrac{1}{2}\lambda^{-2}). \tag{98}$$

Lemma 15 *For each $\lambda > 0$*

$$\mu(q \,||\, \lambda) = \mathbb{E}\left[\left(\frac{B_1^2}{\lambda^2 + B_1^2}\right)^q\right] = \mathbb{E}(|B_1|^{2q})\, h_{-2q}(\lambda) \qquad (q > -\tfrac{1}{2}) \tag{99}$$

where $\mathbb{E}(|B_1|^{2q})$ is given by (93) and h_{-2q} is the Hermite function of index $-2q$, that is $h_0(\lambda) = 1$ and for $q \notin \{0, 1, 2 \ldots\}$

$$h_{-2q}(\lambda) := \frac{1}{2\Gamma(2q)} \sum_{j=0}^{\infty} \Gamma(q + j/2) 2^{q+j/2} \frac{(-\lambda)^j}{j!}. \tag{100}$$

Also,

$$\mu(q \,||\, \lambda) = \mathbb{E}[\exp(-\lambda\sqrt{2\Gamma_q}] \qquad (q > 0) \tag{101}$$

where Γ_q denotes a Gamma random variable with parameter q:

$$\mathbb{P}(\Gamma_q \in dt) = \Gamma(q)^{-1} t^{q-1} e^{-t} dt \qquad (t > 0).$$

Proof. The first equality in (99) is read from (96). The second equality in (99) is the integral representation of the Hermite function provided by Lebedev [35, Problem 10.8.1], and (100) is read from [35, (10.4.3)]. According to another well known integral representation of the Hermite function [35, (10.5.2)], [16, 8.3 (3)], for $q > 0$

$$h_{-2q}(x) = \frac{1}{\Gamma(2q)} \int_0^\infty t^{2q-1} e^{-\frac{1}{2}t^2 - xt} dt = \frac{2^{q-1}}{\Gamma(2q)} \int_0^\infty v^{q-1} e^{-v - x\sqrt{2v}} dv. \tag{102}$$

Formula (101) follows easily from this and (99). $\square$

The identity

$$\mathbb{E}\left[\left(\frac{B_1^2}{\lambda^2+B_1^2}\right)^q\right] = \mathbb{E}[\exp(-\lambda\sqrt{2\Gamma_q}] \quad (q>0), \tag{103}$$

which is implied by the previous proposition, can also be checked by the following argument suggested by Marc Yor. Let X be a positive random variable independent of Γ_q, and let ε with $\varepsilon \stackrel{d}{=} \Gamma_1$ be a standard exponential variable independent of both X and Γ_q. Then by elementary conditioning arguments, for $\theta \geq 0$

$$\mathbb{E}\left[\left(\frac{X}{\theta+X}\right)^q\right] = \mathbb{E}\left[e^{-\theta\Gamma_q/X}\right] = \mathbb{P}(\varepsilon X/\Gamma_q > \theta). \tag{104}$$

Take $X = B_1^2$ and $\theta = \lambda^2$, and use the identity $\varepsilon B_1^2 \stackrel{d}{=} \varepsilon^2/2$, which is a well known probabilistic expression of the gamma duplication formula, to deduce (103) from (104).

The following display identifies $h_\nu(z)$ in the notation of various authors:

$$\begin{aligned} h_\nu(z) &= 2^{-\nu/2}H_\nu(z/\sqrt{2}) = 2^{\nu/2}\Psi(-\nu/2,1/2,z^2/2) && \text{(Lebedev[35])} \\ &= 2^{\nu/2}U(-\nu/2,1/2,z^2/2) && \text{(Abramowitz and Stegun) [1]} \\ &= e^{\frac{1}{4}z^2}U(-\nu-\tfrac{1}{2},z) && \text{(Miller[39])} \\ &= e^{\frac{1}{4}z^2}D_\nu(z) && \text{(Erdelyi [16], Toscano [56])} \end{aligned}$$

The functions $U(a,z)$ and $D_\nu(z)$ are known as *parabolic cylinder functions*, *Weber functions* or *Whittaker functions*. The function $U(a,b,z)$, which is available in *Mathematica* as `HypergeometricU[a,b,z]`, is a *confluent hypergeometric function of the second kind*. Note that $h_n(z)$ defined for $n = 0,1,2,\ldots$ by continuous extension of (100) is the sequence of Hermite polynomials orthogonal with respect to the standard Gaussian density $\varphi(x)$. Also, the function $h_{-1}(x)$ for real x is identified as *Mill's ratio* [26, 33.7]:

$$h_{-1}(x) = \frac{\mathbb{P}(B_1 > x)}{\varphi(x)} = e^{\frac{1}{2}x^2}\int_x^\infty e^{-\frac{1}{2}z^2}dz. \tag{105}$$

For all complex ν and z, the Hermite function satisfies the recursion

$$h_{\nu+1}(z) = zh_\nu(z) - \nu h_{\nu-1}(z), \tag{106}$$

which combined with (105) and $h_0(x) = 1$ yields

$$h_{-2}(x) = 1 - xh_{-1}(x) \tag{107}$$

$$2!h_{-3}(x) = -x + (1+x^2)h_{-1}(x) \tag{108}$$

$$3!h_{-4}(x) = 2 + x^2 - (3x+x^3)h_{-1}(x) \tag{109}$$

and so on. See [51] for further interpretations of the Hermite function in terms of Brownian motion and related stochastic processes.

8.2 Partition probabilities

Recall the notation

$$[\tfrac{1}{2}]_n := \prod_{j=1}^{n}(j-\tfrac{1}{2}) = \frac{\Gamma(\frac{1}{2}+n)}{\Gamma(\frac{1}{2})} = \frac{(2n)!}{2^{2n}n!}.$$

Corollary 16 *The distribution of $\Pi(\lambda)$, a Brownian excursion partition conditioned on $L_1 = \lambda$, is determined by the following EPPF: for $n_1, \ldots, n_k$ with $\sum_{i=1}^k n_i = n$*

$$p_{\frac{1}{2}}(n_1, \ldots, n_k \,||\, \lambda) = 2^{n-k}\lambda^{k-1}h_{k+1-2n}(\lambda)\prod_{i=1}^{k}[\tfrac{1}{2}]_{n_i-1}. \tag{110}$$

Proof. This is read from (66), (99) and (93). □

Formula (110) combined with (14) gives an expression in terms of the Hermite function for the positive integer moments of the sum $S_m(\lambda)$ of mth powers of lengths of excursions of Brownian motion on $[0,1]$ given $L_1 = \lambda$. This formula for $m = 2$ was derived in another way by Janson [25, Theorem 7.4]. There the distribution of $S_2(\lambda)$ appears as the asymptotic distribution, in a suitable limit regime, of the cost of linear probing hashing.

According to (91) and Definition 11, for each $\theta > -\frac{1}{2}$, the EPPF (110) describes the conditional distribution of a $\mathrm{PD}(\frac{1}{2}, \theta)$ partition (Π_n) given $\lim_n K_n/\sqrt{2n} = \lambda$, where K_n is the number of blocks of Π_n. Easily from (110), for each fixed $\lambda > 0$, a sequential description of $(\Pi_n(\lambda), n = 1, 2, \ldots)$ is obtained by replacing the prediction rules (75) and (76) by

$$\mathbb{P}(j\uparrow \,|\, n_1, \cdots, n_k) = (2n_j - 1)\,\frac{h_{k-1-2n}(\lambda)}{h_{k+1-2n}(\lambda)} \quad (1 \le j \le k) \tag{111}$$

$$\mathbb{P}(k+1\uparrow \,|\, n_1, \cdots, n_k) = \frac{\lambda h_{k-2n}(\lambda)}{h_{k+1-2n}(\lambda)}. \tag{112}$$

The addition rule for the EPPF (110) is equivalent to the fact that these transition probabilities sum to 1. As a check, this is implied the recurrence formula (106) for the Hermite function.

Corollary 17 *Let $K_n(\lambda)$ be the number of blocks of $\Pi_n(\lambda)$, where $(\Pi_n(\lambda), n = 1, 2, \ldots)$ is the Brownian excursion partition conditioned on $L_1 = \lambda$. Then $(K_n(\lambda), n = 1, 2, \ldots)$ is a Markov chain with the following inhomogeneous transition probabilities: for $1 \le k \le n$*

$$\mathbb{P}(K_{n+1}(\lambda) = k \,|\, K_n(\lambda) = k) = (2n-k)\,\frac{h_{k-1-2n}(\lambda)}{h_{k+1-2n}(\lambda)} \tag{113}$$

$$\mathbb{P}(K_{n+1}(\lambda) = k+1 \,|\, K_n(\lambda) = k) = \frac{\lambda h_{k-2n}(\lambda)}{h_{k+1-2n}(\lambda)}. \tag{114}$$

Moreover, the distribution of $K_n(\lambda)$ is given by the formula

$$\mathbb{P}(K_n(\lambda) = k) = \frac{(2n-k-1)!\, \lambda^{k-1} h_{k+1-2n}(\lambda)}{(n-k)!(k-1)!2^{n-k}} \qquad (1 \le k \le n). \tag{115}$$

Proof. The Markov property of $(K_n(\lambda), n = 1, 2, \ldots)$ and the transition probabilities (113)–(114) follow easily from (111)–(112). Then (115) follows by induction on n, using the forwards equations implied by the transition probabilities. □

Let K_n denote the number of blocks of Π_n, where (Π_n) is the unconditioned Brownian excursion partition. Then, from the discussion around (90),

$$(K_n(\lambda), n \ge 1) \stackrel{d}{=} (K_n, n \ge 1 \,|\, \lim_n K_n/\sqrt{2n} = \lambda). \tag{116}$$

According to (89), (75) and (76), the sequence $(K_n, n \ge 1)$ is an inhomogeneous Markov chain with transition probabilities

$$\mathbb{P}(K_{n+1} = k \,|\, K_n = k) = \frac{2n-k}{2n} \tag{117}$$

$$\mathbb{P}(K_{n+1} = k+1 \,|\, K_n = k) = \frac{k}{2n} \tag{118}$$

which imply that the unconditional distribution of K_n is given by the formula [46, Corollary 3]

$$\mathbb{P}(K_n = k) = \binom{2n-k-1}{n-1} 2^{k+1-2n} \qquad (1 \le k \le n). \tag{119}$$

Due to (116), for each $\lambda > 0$ the inhomogeneous Markov chain $(K_n(\lambda), n \ge 1)$ has the same co-transition probabilities as $(K_n, n \ge 1)$. From (117), (118) and (119), the co-transition probabilities of $(K_n, n \ge 1)$ are

$$\mathbb{P}(K_n = k \,|\, K_{n+1} = k) = \frac{2(n-k+1)}{2n-k+1} \tag{120}$$

$$\mathbb{P}(K_n = k-1 \,|\, K_{n+1} = k) = \frac{k-1}{2n-k+1}. \tag{121}$$

As a check, the fact that $(K_n(\lambda), n \ge 1)$ has the same co-transition probabilities can be read from (113), (114) and (115). It can be shown that the Markov chains $(K_n(\lambda), n \ge 1)$ for $\lambda \in [0, \infty]$, with definition by weak continuity for $\lambda = 0$ or ∞, are the extreme points of the convex set of all laws of Markov chains with these co-transition probabilities. A generalization of this fact, to $\alpha \in (0,1)$ instead of $\alpha = \frac{1}{2}$, and similar considerations for $\alpha = 0$, yield the second sentence of Theorem 8.

To illustrate the formulas above, according to (9) and (99), or (110) for $n = 2$, given $L_1 = \lambda$, two independent uniform $[0, 1]$ variables fall in the same excursion interval of the Brownian motion with probability

$$p_{\frac{1}{2}}(2 \,||\, \lambda) = \mu(1 \,||\, \lambda) = h_{-2}(\lambda) = 1 - \lambda h_{-1}(\lambda) \tag{122}$$

and in different excursion intervals with probability $\lambda h_{-1}(\lambda)$. According to (110) for $n = 3$, given $L_1 = \lambda$, three independent uniform random points U_1, U_2, U_3 with uniform distribution on $[0, 1]$ fall in the same excursion interval of a Brownian motion or Brownian bridge with probability

$$\mathbb{P}(K_3(\lambda) = 1) = p_{\frac{1}{2}}(3 \,||\, \lambda) = 3h_{-4}(\lambda) = 1 + \tfrac{1}{2}\lambda^2 - (\tfrac{3}{2}\ \lambda + \tfrac{1}{2}\lambda^3)h_{-1}(\lambda) \tag{123}$$

while U_1 and U_2 fall in one excursion interval and U_3 in another with probability

$$\tfrac{1}{3}\,\mathbb{P}(K_3(\lambda) = 2) = p_{\frac{1}{2}}(2, 1 \,||\, \lambda) = \lambda h_{-3}(\lambda) = -\tfrac{1}{2}\lambda^2 + (\tfrac{1}{2}\lambda + \tfrac{1}{2}\lambda^3)h_{-1}(\lambda) \tag{124}$$

and the three points fall in three different excursion intervals with probability

$$\mathbb{P}(K_3(\lambda) = 3) = p_{\frac{1}{2}}(1, 1, 1 \,||\, \lambda) = \lambda^2 h_{-2}(\lambda) = \lambda^2 - \lambda^3 h_{-1}(\lambda). \tag{125}$$

As a check, the sum of expressions for $\mathbb{P}(K_3(\lambda) = k)$ over $k = 1, 2, 3$ reduces to 1. Since

$$\mathbb{P}(K_n(\lambda) = k) = \sum_{n_1 \geq \cdots \geq n_k} \#(n_1, \cdots, n_k) p_{\frac{1}{2}}(n_1, \cdots, n_k \,||\, \lambda) \tag{126}$$

where the sum is over all decreasing sequences of positive integers $(n_1, \cdots, n_k)$ with sum n, and $\#(n_1, \cdots, n_k)$ is the number of distinct partitions of $\mathbb{N}_n$ into k subsets of sizes $(n_1, \cdots, n_k)$, formula (115) amounts to

$$\sum_{n_1 \geq \cdots \geq n_k} \#(n_1, \cdots, n_k) \prod_{i=1}^{k} [\tfrac{1}{2}]_{n_i - 1} = \binom{2n - k - 1}{n - 1} \frac{\Gamma(n)}{\Gamma(k)}\, 2^{2k - 2n} \tag{127}$$

which can be checked as follows. According to (74) and (89), the unconditional EPPF of the Brownian excursion partition Π is

$$p_{\frac{1}{2},0}(n_1, \cdots, n_k) = \frac{\Gamma(k)}{2^{k-1}\Gamma(n)} \prod_{i=1}^{k} [\tfrac{1}{2}]_{n_i - 1} \tag{128}$$

so (127) can be deduced from (128), (119), and the unconditioned form of (126).

8.3 Some identitities

As a consequence of (92) and (99), for all $q > -\frac{1}{2}$ and $\theta > -\frac{1}{2}$ there is the identity

$$\frac{2}{\mathbb{E}(|B_1|^{2\theta})} \int_0^\infty \lambda^{2\theta} \mu(q \,||\, \lambda) \phi(\lambda) d\lambda = \frac{\Gamma(\theta+1)\Gamma(q+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(q+\theta+1)} \tag{129}$$

where the right side is the qth moment of the beta$(\frac{1}{2}, \frac{1}{2}+\theta)$ structural distribution of PD$(\frac{1}{2}, \theta)$, and on the left side this moment is computed by conditioning on L_1. As in (80), for each fixed q this formula provides a Mellin transform which uniquely determines $\mu(q \,||\, \lambda)$ as a function of λ. In view of (129) and (93), the formula (99) for $\mu(q \,||\, \lambda)$ in terms of the Hermite function amounts to the identity

$$2 \int_0^\infty \lambda^{2\theta} h_{-2q}(\lambda) \phi(\lambda) d\lambda = 2^{-\theta-q} \frac{\Gamma(2\theta+1)}{\Gamma(q+\theta+1)}. \tag{130}$$

As checks, since $h_0(x) = 1$ and $h_{-1}(x) = \Phi(x)/\varphi(x)$, the case $q = 0$ is obvious, and the case $q = \frac{1}{2}$ is easily verified since then the left side of (129) equals $(2\theta+1)^{-1}\mathbb{E}(|B_1|^{2\theta+1})$ by integration by parts. Formula (130) can then be verified for $q = m/2$ for all $m = 0, 1, 2, \ldots$, using the recursion (106). Formula (130) was just derived for $q > -\frac{1}{2}$, but both sides are entire functions of q, so the identity holds for all $q \in \mathbb{C}$. Using the series formula (100) and integrating term by term, the substitution $r = \theta + \frac{1}{2}$ allows the identity (130) to be rewritten in the symmetric form

$$\sum_{j=0}^\infty \Gamma\left(q + \frac{j}{2}\right) \Gamma\left(r + \frac{j}{2}\right) \frac{(-2)^j}{j!} = \frac{4\sqrt{\pi}\Gamma(2q)\Gamma(2r)}{\Gamma(q+r+1/2)} \tag{131}$$

where the series is absolutely convergent for real q and r with $q + r + \frac{1}{2} < -1$, and can otherwise be summed by Abel's method provided neither $2q$ nor $2r$ is a non-positive integer. This version of the identity is easily verified using standard identities involving Gauss's hypergeometric function and the gamma function. For $-2q = n$ a positive integer, when h_n is the nth Hermite polynomial

$$h_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} h_{n,k}\, x^{n-2k} \text{ with } h_{n,k} = (-1)^k \binom{n}{2k} \frac{(2k)!}{2^k k!}$$

the identity (130) reduces easily to the following pair of identities of polynomials in θ, which relate the rising and falling factorials $[x]_n := x(x+1)\cdots(x+n-1)$ and $(x)_n := x(x-1)\cdots(x-n+1)$, and which are easily verified directly: for $m = 0, 1, 2 \ldots$

$$\sum_{k=0}^m h_{2m,k} 2^{-k} [\theta + \tfrac{1}{2}]_{m-k} = (\theta)_m$$

and

$$\sum_{k=0}^m h_{2m+1,k} 2^{-k} [\theta + 1]_{m-k} = (\theta - \tfrac{1}{2})_m.$$

Thus the coefficients of the Hermite polynomials are related to some instances of generalized Stirling numbers [22, 48].

Acknowledgment

Thanks to Grégory Miermont for careful reading of a draft of this paper.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover, New York, 1965.

[2] D. Aldous and J. Pitman. Brownian bridge asymptotics for random mappings. *Random Structures and Algorithms*, 5:487–512, 1994.

[3] D.J. Aldous and J. Pitman. The standard additive coalescent. *Ann. Probab.*, 26:1703–1726, 1998.

[4] R. A. Arratia, A. D. Barbour, and S. Tavaré. Logarithmic combinatorial structures: a probabilistic approach. Book in preparation. Available via `www-hto.usc.edu/books/tavare/ABT/index.html`, 2001.

[5] J. Bertoin. A fragmentation process connected to Brownian motion. *Probab. Theory Related Fields*, 117(2):289–301, 2000.

[6] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

[7] D. Blackwell and J.B. MacQueen. Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, 1:353–355, 1973.

[8] A. Brix. Generalized gamma measures and shot-noise Cox processes. *Adv. in Appl. Probab.*, 31(4):929–953, 1999.

[9] P. Chassaing and S. Janson. A Vervaat-like path transformation for the reflected Brownian bridge conditioned on its local time at 0. *Ann. Probab.*, 29(4):1755–1779, 2001.

[10] P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms*, 21:76–119, 2002.

[11] B. Derrida. Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B (3)*, 24(5):2613–2626, 1981.

[12] B. Derrida. Non-self-averaging effects in sums of random variables, spin glasses, random maps and random walks. In M. Fannes, C. Maes, and A. Verbeure, editors, *On Three Levels: Micro-, Meso-, and Macro-Approaches in Physics*, NATO ASI Series, pages 125–137. Plenum Press, New York and London, 1994.

[13] B. Derrida. From random walks to spin glasses. *Phys. D*, 107(2-4):186–198, 1997. Landscape paradigms in physics and biology (Los Alamos, NM, 1996).

[14] G. Doetsch. *Theorie und Anwendung der Laplace-Transformation*. Berlin, 1937.

[15] S. Engen. *Stochastic Abundance Models with Emphasis on Biological Communities and Species Diversity*. Chapman and Hall Ltd., 1978.

[16] A. Erdélyi et al. *Higher Transcendental Functions*, volume II of *Bateman Manuscript Project*. McGraw-Hill, New York, 1953.

[17] W. J. Ewens and S. Tavaré. The Ewens sampling formula. In N. S. Johnson, S. Kotz, and N. Balakrishnan, editors, *Multivariate Discrete Distributions*. Wiley, New York, 1995.

[18] W.J. Ewens. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.*, 3:87 – 112, 1972.

[19] M. Grote and T. P. Speed. Approximate Ewens formulae for symmetric overdominance selection. *Annals of Applied Probability*, 12: 637 – 663, 2002.

[20] B. Hansen and J. Pitman. Prediction rules and exchangeable sequences related to species sampling. *Stat. and Prob. Letters*, 46(520):251–256, 2000.

[21] F. M. Hoppe. The sampling theory of neutral alleles and an urn model in population genetics. *Journal of Mathematical Biology*, 25:123 – 159, 1987.

[22] L. C. Hsu and P. J.-S. Shiue. A unified approach to generalized Stirling numbers. *Adv. in Appl. Math.*, 20(3):366–384, 1998.

[23] L. F. James. Poisson process partition calculus with applications to exchangeable models and Bayesian nonparametrics. arXiv:math.PR/0205093, 2002.

[24] L.F. James. Bayesian calculus for gamma processes with applications to semiparametric intensity models. *Sankhyā*, 2002. to appear.

[25] S. Janson. Asymptotic distribution for the cost of linear probing hashing. *Random Structures Algorithms*, 19(3-4):438–471, 2001. Analysis of algorithms (Krynica Morska, 2000).

[26] N. L. Johnson and S. Kotz. *Continuous Univariate Distributions, volume 2*. Wiley, 1970.

[27] S. Karlin. Central limit theorems for certain infinite urn schemes. *J. Math. Mech.*, 17:373–401, 1967.

[28] S. Kerov. Coherent random allocations and the Ewens-Pitman formula. PDMI Preprint, Steklov Math. Institute, St. Petersburg, 1995.

[29] J. F. C. Kingman. Random discrete distributions. *J. Roy. Statist. Soc. B*, 37:1–22, 1975.

[30] J. F. C. Kingman. The representation of partition structures. *J. London Math. Soc.*, 18:374–380, 1978.

[31] J. F. C. Kingman. The coalescent. *Stochastic Processes and their Applications*, 13:235–248, 1982.

[32] J. F. C. Kingman. Exchangeability and the evolution of large populations. In G. Koch and F. Spizzichino, editors, *Exchangeability in probability and statistics (Rome, 1981)*, pages 97–112. North-Holland, Amsterdam, 1982.

[33] J. F. C. Kingman. *Poisson Processes*. Clarendon Press, Oxford, 1993.

[34] U. Küchler and M. Sørensen. *Exponential families of stochastic processes*. Springer-Verlag, New York, 1997.

[35] N. N. Lebedev. *Special Functions and their Applications*. Prentice-Hall, Englewood Cliffs, N.J., 1965.

[36] P. Lévy. Sur certains processus stochastiques homogènes. *Compositio Math.*, 7:283–339, 1939.

[37] J. W. McCloskey. A model for the distribution of individuals by species in an environment. Ph. D. thesis, Michigan State University, 1965.

[38] G. Miermont and J. Schweinsberg. Self-similar fragmentations and stable subordinators. Technical Report 726, 2001. Prépublication du Laboratoire de Probabilités et modèles aléatoires, Universite Paris VI. Available via `http://www.proba.jussieu.fr`.

[39] J. C. P. Miller. *Tables of Weber parabolic cylinder functions*. Her Majesty's Stationery Office, London, 1955.

[40] M. Perman. Order statistics for jumps of normalized subordinators. *Stoch. Proc. Appl.*, 46:267–281, 1993.

[41] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. *Probab. Th. Rel. Fields*, 92:21–39, 1992.

[42] J. Pitman. Poisson-Kingman partitions. Preprint available via `www.stat.berkeley.edu/users/pitman`, 1995.

[43] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probab. Th. Rel. Fields*, 102:145–158, 1995.

[44] J. Pitman. Random discrete distributions invariant under size-biased permutation. *Adv. Appl. Prob.*, 28:525–539, 1996.

[45] J. Pitman. Some developments of the Blackwell-MacQueen urn scheme. In T.S. Ferguson et al., editor, *Statistics, Probability and Game Theory; Papers in honor of David Blackwell*, volume 30 of *Lecture Notes-Monograph Series*, pages 245–267. Institute of Mathematical Statistics, Hayward, California, 1996.

[46] J. Pitman. Partition structures derived from Brownian motion and stable subordinators. *Bernoulli*, 3:79–96, 1997.

[47] J. Pitman. Brownian motion, bridge, excursion and meander characterized by sampling at independent uniform times. *Electron. J. Probab.*, 4:Paper 11, 1–33, 1999.

[48] J. Pitman. Combinatorial Stochastic Processes. Technical Report 621, Dept. Statistics, U.C. Berkeley, 2002. Lecture notes for St. Flour course, July 2002. Available via `www.stat.berkeley.edu`.

[49] J. Pitman and M. Yor. Arcsine laws and interval partitions derived from a stable subordinator. *Proc. London Math. Soc. (3)*, 65:326–356, 1992.

[50] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, 25:855–900, 1997.

[51] J. Pitman and M. Yor. Brownian interpretations of Mill's ratio, Hermite and parabolic cylinder functions. In preparation, 2002.

[52] H. Pollard. The representation of $e^{-x^{\lambda}}$ as a Laplace integral. *Bull. Amer. Math. Soc.*, 52:908–910, 1946.

[53] D. Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108:225–239, 1987.

[54] V. Seshadri. *The inverse Gaussian distribution.* The Clarendon Press, Oxford University Press, New York, 1993.

[55] M. Talagrand. *Spin glasses: a challenge to mathematicians.* Springer, 2001. Book in preparation.

[56] L. Toscano. Sulle funzioni del cilindro parabolico. *Matematiche (Catania)*, 26:104–126 (1972), 1971.

[57] S.L. Zabell. The continuum of inductive methods revisited. In J. Earman and J. D. Norton, editors, *The Cosmos of Science*, Pittsburgh-Konstanz Series in the Philosophy and History of Science, pages 351–385. University of Pittsburgh Press/Universitätsverlag Konstanz, 1997.

Jim Pitman, Department of Statistics, University of California, Berkeley
`pitman@stat.berkeley.edu`