The Annals of Probability
1997, Vol. 25, No. 2, 855–900

THE TWO-PARAMETER POISSON–DIRICHLET DISTRIBUTION DERIVED FROM A STABLE SUBORDINATOR[1]

By Jim Pitman and Marc Yor

University of California, Berkeley

The two-parameter Poisson–Dirichlet distribution, denoted PD(α, θ), is a probability distribution on the set of decreasing positive sequences with sum 1. The usual Poisson–Dirichlet distribution with a single parameter θ, introduced by Kingman, is PD$(0, \theta)$. Known properties of PD$(0, \theta)$, including the Markov chain description due to Vershik, Shmidt and Ignatov, are generalized to the two-parameter case. The size-biased random permutation of PD(α, θ) is a simple residual allocation model proposed by Engen in the context of species diversity, and rediscovered by Perman and the authors in the study of excursions of Brownian motion and Bessel processes. For $0 < \alpha < 1$, PD$(\alpha, 0)$ is the asymptotic distribution of ranked lengths of excursions of a Markov chain away from a state whose recurrence time distribution is in the domain of attraction of a stable law of index α. Formulae in this case trace back to work of Darling, Lamperti and Wendel in the 1950s and 1960s. The distribution of ranked lengths of excursions of a one-dimensional Brownian motion is PD$(1/2, 0)$, and the corresponding distribution for a Brownian bridge is PD$(1/2, 1/2)$. The PD$(\alpha, 0)$ and PD(α, α) distributions admit a similar interpretation in terms of the ranked lengths of excursions of a semistable Markov process whose zero set is the range of a stable subordinator of index α.

CONTENTS

Received August 1995; revised July 1996.
[1]Research supported in part by NSF Grant DMS-94-04345.
AMS 1991 *subject classifications*. Primary 60J30; secondary 60G57, 60E99.
Key words and phrases. Zero set, semistable Markov process, local time, ranked lengths of excursions, Poisson point process.

1. Introduction. The subject of this paper is a two-parameter family of probability distributions for a sequence of random variables

$$(V_n) = (V_1, V_2, \ldots) \quad \text{with } V_1 > V_2 > \cdots > 0 \quad \text{and } \sum_n V_n = 1 \text{ a.s.} \tag{1}$$

This family extends the one-parameter family of Poisson–Dirichlet distributions, introduced by Kingman [38] and denoted here by PD(0, θ), $\theta > 0$, which arises from the study of asymptotic distributions of random ranked relative frequencies in a variety of contexts including number theory [6, 65], combinatorics [66, 1, 27], Bayesian statistics [22] and population genetics [72, 20]. Study of an enlarged family, involving another parameter α with $0 \le \alpha < 1$, is motivated by parallels between PD(0, θ) and the asymptotic distributions of ranked relative lengths of intervals derived in renewal theory from lifetime distributions in the domain of attraction of a stable law of index α [42, 74]. As explained in Section 1.2, this family of asymptotic distributions for (V_n) as in (1), denoted here by PD(α, 0), $0 < \alpha < 1$, can be interpreted in terms of ranked lengths of excursion intervals between zeros of B, where B is Brownian motion for $\alpha = 1/2$, or a recurrent Bessel process of dimension $2 - 2\alpha$ for $0 < \alpha < 1$. By a change of measure relative to PD(α, 0), with a density depending on θ described in Proposition 14, we can define PD(α, θ) for arbitrary $0 < \alpha < 1$ and $\theta > -\alpha$, then recover Kingman's Poisson-Dirichlet distribution PD(0, θ) for $\theta > 0$ as the weak limit of PD(α, θ) as $\alpha \downarrow 0$. We prefer, however, to present a unified definition of PD(α, θ) as follows.

1.1. *The size-biased permutation of* PD(α, θ). The following definition originates from the application of random discrete distributions to model the division of a large population into a large number of possible species or types. A ranked sequence of random frequencies (V_n) as in (1) represents the structure of an idealized infinite population which has been randomly partitioned into various species. Then V_n represents the proportion of the population that belongs to the nth most common species. See [17, 38, 20, 56] for background and further references to such applications. The *size-biased permutation* of (V_n) is the sequence of proportions of species in their order of appearance in a process of random sampling from the population. This notion is made precise as follows. For (V_n) as in (1), call a random variable $\tilde{V}_1$ a *size-biased pick* from (V_n) if

$$P(\tilde{V}_1 = V_n | V_1, V_2, \dots) = V_n \qquad (n = 1, 2, \dots). \tag{2}$$

Here $\tilde{V}_1$ may be already defined on the same probability space as (V_n) or constructed by additional randomization on an enlarged probability space. Call $(\tilde{V}_1, \tilde{V}_2, \dots)$ a *size-biased permutation* of (V_n) if $\tilde{V}_1$ is a size-biased pick from (V_n), and for each $n = 1, 2, \dots$ and $j = 1, 2, \dots$,

$$P(\tilde{V}_{n+1} = V_j | \tilde{V}_1, \dots, \tilde{V}_n; V_1, V_2, \dots) = \frac{V_j 1(V_j \neq \tilde{V}_i \text{ for all } 1 \leq i \leq n)}{(1 - \tilde{V}_1 - \cdots - \tilde{V}_n)}.$$

Following Engen [17] and Perman, Pitman and Yor [51], we make the following definition in terms of independent beta random variables. See also Appendixes A.1 and A.2 for further motivation. Recall that for $a > 0$, $b > 0$, the beta(a, b) *distribution* on $(0, 1)$ has density

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \qquad (0 < x < 1). \tag{3}$$

DEFINITION 1. For $0 \leq \alpha < 1$ and $\theta > -\alpha$, suppose that a probability $P_{\alpha,\theta}$ governs independent random variables $\tilde{Y}_n$ such that $\tilde{Y}_n$ has beta$(1-\alpha, \theta+n\alpha)$ distribution. Let

$$\tilde{V}_1 = \tilde{Y}_1, \qquad \tilde{V}_n = (1-\tilde{Y}_1)\cdots(1-\tilde{Y}_{n-1})\tilde{Y}_n \qquad (n \geq 2) \tag{4}$$

and let $V_1 \geq V_2 \geq \cdots$ be the ranked values of the $\tilde{V}_n$. Define the *Poisson–Dirichlet distribution with parameters* (α, θ), abbreviated PD(α, θ), to be the $P_{\alpha,\theta}$ distribution of (V_n).

Results of [51] show that this definition of PD(α, θ) agrees with the previous descriptions of PD$(0, \theta)$ and PD$(\alpha, 0)$ and yield the following result.

PROPOSITION 2 [48, 51, 56]. *Under $P_{\alpha,\theta}$ governing $(\tilde{Y}_n)$, $(\tilde{V}_n)$ and (V_n) as in Definition* 1, *the sequence (V_n) is such that $V_1 > V_2 > \cdots > 0$ and $\sum_n V_n = 1$ almost surely, and $(\tilde{V}_n)$ is a size-biased permutation of (V_n).*

To put the result of Proposition 2 another way, suppose that (V_n) is any sequence of random variables with PD(α, θ) distribution for some $0 \leq \alpha < 1$,

$\theta > -\alpha$, that $(\tilde{V}_n)$ is a size-biased permutation of (V_n) and let

$$\tilde{Y}_n = \tilde{V}_n/(\tilde{V}_n + \tilde{V}_{n+1} + \cdots). \tag{5}$$

Then these three sequences (V_n), $(\tilde{V}_n)$ and $(\tilde{Y}_n)$ have the same joint distribution as those in Definition 1. In particular Proposition 2 implies the following corollary.

COROLLARY 3 [48, 17, 51, 56]. *For* $0 \le \alpha < 1$ *and* $\theta > -\alpha$, *if* $\tilde{V}_1$ *is a size-biased pick from* (V_n) *with* PD(α, θ) *distribution, then* $\tilde{V}_1$ *has* beta$(1-\alpha, \theta+\alpha)$ *distribution.*

As a consequence of Corollary 3,

$$\begin{aligned} E_{\alpha,\theta} \sum_{n=1}^{\infty} f(V_n) &= E_{\alpha,\theta}\left[\frac{f(\tilde{V}_1)}{\tilde{V}_1}\right] \\ &= \frac{\Gamma(\theta+1)}{\Gamma(\theta+\alpha)\Gamma(1-\alpha)} \int_0^1 du\, f(u) \frac{(1-u)^{\alpha+\theta-1}}{u^{\alpha+1}}, \end{aligned} \tag{6}$$

where we revert to the setting of Definition 1, with $E_{\alpha,\theta}$ denoting expectation with respect to the probability distribution $P_{\alpha,\theta}$.

The result of Proposition 2 for $\alpha = 0$ is due to McCloskey [48]. Ewens [20] called the $P_{0,\theta}$ distribution of $(\tilde{V}_n)$ defined by (4) the GEM distribution, after Griffiths, Engen and McCloskey. Engen [17] considered also the residual allocation model (4) for $(\tilde{V}_n)$ for $0 \le \alpha < 1$ and $\theta > 0$, and he established Corollary 3 for this range of parameters. The particular choice of beta distributions for $\tilde{Y}_n$ in Definition 1, and the consequent parameter set $\{0 \le \alpha < 1, \theta > -\alpha\}$ for the two-parameter Poisson–Dirichlet distribution, is dictated by the following result, which generalizes a well known characterization of PD$(0, \theta)$ due to McCloskey [48].

PROPOSITION 4 [56]. *For* (V_n) *with* $V_1 > V_2 > \cdots > 0$ *and* $\sum_n V_n = 1$ *almost surely, a size-biased random permutation* $(\tilde{V}_n)$ *of* (V_n) *admits the expression* (4) *for a sequence of independent random variables* $(\tilde{Y}_n)$ *iff the distribution of the* $\tilde{Y}_n$ *is of the form assumed in Definition 1, that is, iff* (V_n) *has* PD(α, θ) *distribution for some* $0 \le \alpha < 1$ *and* $\theta > -\alpha$.

1.2. *Interval lengths derived from a subordinator.* Following Lamperti [42, 43], Wendel [74], Kingman [38] and Perman, Pitman and Yor [50, 51, 59], consider the sequence

$$V_1(T) \ge V_2(T) \ge \cdots \ge 0 \tag{7}$$

of ranked lengths of component intervals of the set $[0, T]\backslash Z$, where Z is a random closed subset of $[0, \infty)$ with Lebesgue measure 0, and T is a strictly positive random time. Suppose Z is the closure of the range of a *subordinator* $(\tau_s, s \geq 0)$, that is, an increasing process with stationary independent increments. Assume that (τ_s) has no drift component, so

$$E[\exp(-\lambda\tau_s)] = \exp\left(-s\int_0^\infty (1-\exp(-\lambda x))\Lambda(dx)\right), \tag{8}$$

where the Lévy measure Λ on $(0,\infty)$ is the intensity measure for the Poisson point process of jumps $(\tau_s - \tau_{s-}, s \geq 0)$. Call (τ_s) a *gamma subordinator* if $\Lambda(dx) = x^{-1}e^{-x}\,dx$, $x > 0$, that is, if τ_s has the *gamma*(s) *distribution*

$$P(\tau_s \in dx) = \Gamma(s)^{-1}x^{s-1}e^{-x}\,dx \qquad (x > 0) \tag{9}$$

for each $s > 0$. There is the following well known representation of PD$(0,\theta)$.

PROPOSITION 5 [48, 22, 38]. *If (τ_s) is a gamma subordinator, then for every $\theta > 0$ the sequence*

$$\left(\frac{V_1(\tau_\theta)}{\tau_\theta}, \frac{V_2(\tau_\theta)}{\tau_\theta}, \ldots\right) \text{ has PD}(0,\theta) \text{ distribution} \tag{10}$$

and is independent of τ_θ.

Let $0 < \alpha < 1$. Call (τ_s) a *stable (α) subordinator* if $\Lambda = \Lambda_\alpha$, where

$$\Lambda_\alpha(x,\infty) = Cx^{-\alpha} \qquad (x > 0) \tag{11}$$

for some constant $C > 0$. That is, from (8), for $\lambda > 0$,

$$E[\exp(-\lambda\tau_s)] = \exp(-sK\lambda^\alpha), \quad \text{where } K = C\Gamma(1-\alpha). \tag{12}$$

The following companion of Proposition 5 plays a key role in this paper.

PROPOSITION 6 [51, 59]. *If (τ_s) is a stable (α) subordinator for some $0 < \alpha < 1$ then for every $s > 0$,*

$$\left(\frac{V_1(\tau_s)}{\tau_s}, \frac{V_2(\tau_s)}{\tau_s}, \ldots\right) \text{ has PD}(\alpha,0) \text{ distribution} \tag{13}$$

and also for every fixed $t > 0$,

$$\left(\frac{V_1(t)}{t}, \frac{V_2(t)}{t}, \ldots\right) \text{ has PD}(\alpha,0) \text{ distribution.} \tag{14}$$

The equality in distribution of the two sequences displayed in (13) and (14) was established in [59], while the connection with Definition 1 was made in [51]. See Section 8.2 of this paper for a characterization of the laws of the sequences displayed in (13) and (14) for a more general subordinator (τ_s), and [5, 59, 71, 60] regarding the relation between description of PD$(\alpha, 0)$ in Proposition 6 and the generalized arcsine laws of Lamperti [41].

In contrast to Proposition 5, the random variable τ_s is not independent of the PD$(\alpha, 0)$ distributed sequence displayed in (13). On the contrary, results of [38, 59] reviewed in Section 2 show that the random variable τ_s is almost surely equal to a measurable function of this sequence. Results of Perman [50] describe the family of conditional distributions of the sequence in (13) given $\tau_s = t$ for $t > 0$; see Section 8.1. A result of [51] reviewed in Section 3 shows that for $0 < \alpha < 1$ and $\theta > -\alpha$ the distribution PD(α, θ) is obtained by mixing these conditional distributions derived from the stable (α) subordinator (τ_s) according to the probability measure with density proportional to $t^{-\theta}$ relative to $P(\tau_s \in dt)$.

Since the zero set Z of a standard one-dimensional Brownian motion B is the closure of the range of a stable (1/2) subordinator [46], (14) shows that PD$(1/2, 0)$ is the distribution of the ranked lengths of the excursions of B away from 0 during the time interval $[0, 1]$. Note that these excursion lengths include the length $1 - G_1$ of the final *meander interval*, where

$$G_t = \sup([0, t) \cap Z) = \sup\{s: s < t, B_s = 0\}. \tag{15}$$

Similarly, PD$(\alpha, 0)$ can be interpreted in terms of the ranked lengths of excursion intervals if the Brownian motion B is replaced by a suitable semistable Markov process [44], for example, a Bessel process of dimension $\delta = 2 - 2\alpha$ [43, 49] or, for $0 < \alpha < 1/2$, a stable Lévy process of index $1/(1 - \alpha)$ [23].

The PD(α, α) distribution arises naturally as the distribution of ranked lengths of excursions of a semistable Markov bridge derived from a Markov process whose zero set is the range of a stable (α) subordinator [74, 59, 51]. It is well known that such a bridge can be derived from the unconditioned process on interval $[0, G_t]$ by appropriate scaling. So as a companion to (13) and (14), in the same setting we have for each fixed $t > 0$,

$$\left(\frac{V_1(G_t)}{G_t}, \frac{V_2(G_t)}{G_t}, \dots\right) \text{ has PD}(\alpha, \alpha) \text{ distribution} \tag{16}$$

independently of G_t. In particular, we note the following proposition:

PROPOSITION 7 [51, 59]. *If V_n is the length of the nth longest excursion of B away from 0 over the time interval $[0, 1]$, then*

$$(V_n) \textit{ has } \text{PD}(1/2, 0) \textit{ distribution if } B \textit{ is Brownian motion;} \tag{17}$$

$$(V_n) \textit{ has } \text{PD}(1/2, 1/2) \textit{ distribution if } B \textit{ is Brownian bridge.} \tag{18}$$

Stepanov [64] encountered asymptotics involving PD$(1/2, 1/2)$ in the study of the asymptotic distribution of the sizes of tree components in a random mapping. The connection with the Brownian bridge in this setting is explained in Aldous and Pitman [3]. See [58, 18] for recent developments in this vein.

The PD$(\alpha, 0)$ distribution also arises as the asymptotic distribution of

$$\left(\frac{V_1(T)}{T}, \frac{V_2(T)}{T}, \dots\right) \tag{19}$$

either for nonrandom T as $T \to \infty$, or for $T = \tau_s$ as $s \to \infty$, for any subordinator (τ_s) such that $\Lambda(x, \infty) = x^{-\alpha}L(x)$ as $x \to \infty$ for a slowly varying function $L(x)$. Similarly, PD$(\alpha, 0)$ is the asymptotic distribution as $n \to \infty$ of

$$\left(\frac{X_{(n,1)}}{S_n}, \frac{X_{(n,2)}}{S_n}, \dots, \frac{X_{(n,n)}}{S_n}\right) \tag{20}$$

for $X_{(n,1)} \geq X_{(n,2)} \geq \cdots \geq X_{(n,n)}$ the order statistics of i.i.d. positive random variables $X_1, \dots, X_n$ with sum S_n, assuming $P(X_i \geq x) = x^{-\alpha}L(x)$ as $x \to \infty$. Related results have been studied by many authors: see, for instance, [12, 4, 42, 31, 32, 61, 76]. Many limit distributions found in these papers are the exact distributions of appropriate functions of a PD$(\alpha, 0)$ sequence. For instance, Darling [12] found the characteristic function of the limiting distribution of $S_n/X_{(n,1)}$ in (20). This is the characteristic function of $1/V_1$ for a PD$(\alpha, 0)$ sequence (V_n). Lamperti [42] derived the corresponding Laplace transform, given by (38) of this paper with $n = 1$, from the asymptotic distribution as $n \to \infty$ of the maximum up to time n of the age process derived from a discrete renewal process with lifetime distribution in the domain of attraction of a stable law of index α. That the same transform appears in both Darling's and Lamperti's works amounts to the equality in distribution of the first components in (13) and (14). The equality in distribution of the first n components in (13) and (14) can be interpreted similarly as an asymptotic result in renewal theory.

1.3. *Organization of the paper.* We develop various results for PD(α, θ) in the general two-parameter case. Most of these results were previously known in either of the special cases $\alpha = 0$ or $\theta = 0$. Many results acquire their simplest form for PD$(\alpha, 0)$ with $0 < \alpha < 1$. These results for PD$(\alpha, 0)$ are presented in Section 2, followed by results for PD(α, θ) in Section 3. These two sections will also serve as a guide to the rest of the paper, which contains proofs of the results in Sections 2 and 3, and various further developments.

2. Main results for PD($\boldsymbol{\alpha}$, 0).

Results stated in this section are proved in Section 4.

PROPOSITION 8. *Suppose (V_n) has* PD$(\alpha, 0)$ *distribution for some $0 < \alpha < 1$. Let*

$$R_n = \frac{V_{n+1}}{V_n}. \tag{21}$$

Then R_n has beta$(n\alpha, 1)$ *distribution, that is,*

$$P(R_n \leq r) = r^{n\alpha} \qquad (0 \leq r \leq 1) \tag{22}$$

and the R_n are mutually independent.

Since (V_n) can be recovered from (R_n) as

$$V_1 = \frac{1}{1 + R_1 + R_1R_2 + R_1R_2R_3 + \cdots}; \tag{23}$$

$$V_{n+1} = V_1 R_1 R_2 \cdots R_n \qquad (n \geq 1),$$

the following simple construction of PD$(\alpha, 0)$is an immediate corollary of Proposition 8.

COROLLARY 9. *Suppose (R_n) is a sequence of independent random variables such that R_n has* beta$(n\alpha, 1)$ *distribution, for some* $0 < \alpha < 1$. *Then (V_n) defined by* (23) *has* PD$(\alpha, 0)$ *distribution.*

The next proposition summarizes and sharpens some results from [38, 59]. The abbreviation "PRM Λ" will be used for "Poisson random measure with intensity measure Λ."

PROPOSITION 10. *Suppose (V_n) has* PD$(\alpha, 0)$ *distribution for some* $0 < \alpha < 1$.

(i) *The limit*

$$L := \lim_{n\to\infty} n V_n^{\alpha} \tag{24}$$

exists both almost surely and in pth mean for all $p \geq 1$.

(ii) *Let*

$$\Sigma := (L/C)^{-1/\alpha}, \qquad \Delta_n := V_n \Sigma. \tag{25}$$

Then Σ has the same stable (α) distribution as τ_1 in (12), *the Δ_n are the ranked points of a* PRM Λ_α *on* $(0, \infty)$, *where $\Lambda_\alpha(x, \infty) = Cx^{-\alpha}$ for $x > 0$, and (V_n) may be represented as*

$$V_n = \Delta_n/\Sigma \quad \text{where } \Sigma = \sum_n \Delta_n, \tag{26}$$

(iii) *Let*

$$X_n := \Lambda_\alpha(\Delta_n, \infty) = C\Delta_n^{-\alpha} = LV_n^{-\alpha}. \tag{27}$$

Then the $X_1 < X_2 < \cdots$ are the points of a PRM (dx) *on* $(0, \infty)$; *that is,*

$$X_n = \varepsilon_1 + \cdots + \varepsilon_n, \tag{28}$$

where the ε_i are independent standard exponential variables and (V_n) may be represented in terms of (X_n) as

$$V_n = \frac{X_n^{-1/\alpha}}{\sum_m X_m^{-1/\alpha}}. \tag{29}$$

Note how in the representation (26), which is a variation of (13), the PD$(\alpha, 0)$ distributed sequence (V_n) is not independent of the sum Σ of the

Poisson points Δ_n. On the contrary, Σ and hence all the Δ_n are recovered as functions of (V_n) via (24) and (25). Compare with the corresponding variation of (10): if the Δ_n are the ranked points of a PRM Λ for $\Lambda(dx) = \theta x^{-1} e^{-x}\,dx$, then (V_n) defined by (26) has PD$(0,\theta)$ distribution independent of Σ. This independence is characteristic of the gamma Lévy measure, due to Lukacs' characterization of the gamma distribution [47] and Kallenberg's representation of the subordinator [35].

Recall that in the setting of Section 1.2, $V_n = V_n(1)$ is the nth longest subinterval in the complement of $[0,1]\backslash Z$, where Z is the zero set of a semistable Markov process X, and L is a multiple of the local time of X at zero up to time 1. So we call the random variable L introduced in (24) the *local time* derived from (V_n). See [60] for further discussion of results with $V_n = V_n(T)/T$ for suitable random T. The distribution of $L = C\Sigma^{-\alpha}$ is determined by its moments

$$E(L^p) = C^p E(\Sigma^{-\alpha p}) = \frac{\Gamma(p+1)}{\Gamma(p\alpha+1)}\Gamma(1-\alpha)^{-p} \qquad (p > -1). \tag{30}$$

So $\Gamma(1-\alpha)L$ has the Mittag–Leffler (α) distribution [23, 49, 7, 51]. The joint distribution of L and $V_1,\dots,V_n$ can be read from that of Σ and $V_1,\dots,V_n$, which is described in Proposition 47. In formula (29), which serves to construct a PD$(\alpha,0)$ sequence (V_n) from a sequence of independent standard exponential variables (ε_n), the denominator has a stable(α) distribution. This method of constructing a random variable with an infinitely divisible distribution from the ranked jumps of its Poisson representation, originally due to Lévy, has been exploited in several contexts [70, 45].

The next proposition exposes some results underlying the following formula (38) for the Laplace transform of $1/V_n$. This formula was obtained in different settings by Darling [12] and Lamperti [42] for $n=1$ and Wendel [74] for $n = 2,3,\dots$. See also Horowitz [32], Kingman [38] and Resnick [61].

PROPOSITION 11. *Suppose (V_n) has* PD$(\alpha,0)$ *distribution for some* $0<\alpha<1$. *Let $A_0 = 0$ and for $n = 1,2,\dots$ define random variables A_n and Σ_n by*

$$A_n := \frac{V_1+V_2+\cdots+V_n}{V_{n+1}} = \frac{1}{R_n} + \frac{1}{R_nR_{n-1}} + \cdots + \frac{1}{R_nR_{n-1}\cdots R_1}, \tag{31}$$

$$\begin{aligned}\Sigma_n &:= \frac{V_{n+1}+V_{n+2}+\cdots}{V_n}\\ &= \frac{1-V_1-\cdots-V_n}{V_n} = R_n + R_nR_{n+1} + R_nR_{n+1}R_{n+2}+\cdots.\end{aligned} \tag{32}$$

where $R_n = V_{n+1}/V_n$ as in Proposition 8. For $\lambda \ge 0$ let

$$\phi_\alpha(\lambda) := \alpha\int_1^\infty dx\, e^{-\lambda x}x^{-\alpha-1}, \tag{33}$$

$$\psi_\alpha(\lambda) := 1 + \alpha\int_0^1 dx\,(1-e^{-\lambda x})x^{-\alpha-1} = \Gamma(1-\alpha)\lambda^\alpha + \phi_\alpha(\lambda) \tag{34}$$

Then

$$\frac{1}{V_n} = 1 + A_{n-1} + \Sigma_n, \tag{35}$$

where:

(i) A_{n-1} *is distributed as the sum of* $n-1$ *independent copies of* A_1, *with*

$$E[\exp(-\lambda A_{n-1})] = \phi_\alpha(\lambda)^{n-1}; \tag{36}$$

(ii) Σ_n *is distributed as the sum of* n *independent copies of* Σ_1 *with*

$$E[\exp(-\lambda \Sigma_n)] = \psi_\alpha(\lambda)^{-n}; \tag{37}$$

(iii) A_{n-1} *and* Σ_n *are independent.*

COROLLARY 12 [12, 42, 74]. *If* (V_n) *has* PD$(\alpha, 0)$ *distribution, then the distribution of* V_n *is determined by the Laplace transform*

$$E[\exp(-\lambda/V_n)] = \exp(-\lambda)\phi_\alpha(\lambda)^{n-1}\psi_\alpha(\lambda)^{-n}. \tag{38}$$

For $V_n = V_n(1)$ derived from the interval lengths $V_n(t)$ generated by the range of a stable (α) subordinator, Wendel [74] obtained (38) by considering the random times

$$H_n := \inf\{t\colon V_n(t) = 1\} \tag{39}$$

for $n = 1, 2, \ldots,$ and using the identity in distribution

$$V_n =_d 1/H_n, \tag{40}$$

which follows by scaling from the equality of events $(H_n > t) = (V_n(t) < 1)$. While both (H_n^{-1}) and (V_n) are decreasing random sequences and (H_n^{-1}) has the same one-dimensional distributions as (V_n), this identity does not extend even to two-dimensional distributions, due to the fact that $\sum_n V_n = 1$ while there is no such constraint on $\sum_n H_n^{-1}$. However, comparison of Wendel's argument with our derivation of Proposition 11 reveals a remarkable extension of the identity in distribution (40).

PROPOSITION 13. *For each* $n = 1, 2, \ldots,$

$$\left(\frac{V_1(H_n)}{H_n}, \frac{V_2(H_n)}{H_n}, \ldots\right) \text{ has PD}(\alpha, 0) \text{ distribution.} \tag{41}$$

See also [60] for some generalizations of Propositions 6 and 13.

Several authors have studied questions related to the a.s. limiting behavior of $V_n(t)$ as $t \to \infty$ for $V_n(t)$ derived from the range of a stable subordinator. See, for example, Chung and Erdös [10], Csaki, Erdös and Revesz [11]. See Hu and Shi [33] for a number of refinements obtained using results of this paper.

3. Main results for PD ($\boldsymbol{\alpha}$, $\boldsymbol{\theta}$).

Results stated in this section are proved in Section 5 except where otherwise indicated. For $0 \le \alpha < 1$ and $\theta > -\alpha$ let $E_{\alpha,\theta}$ denote expectation with respect to the probability $P_{\alpha,\theta}$ governing $(\tilde{V}_n)$ and (V_n) as in Definition 1. So the $P_{\alpha,\theta}$ distribution of (V_n) is PD(α, θ).

3.1. *Change of measure formulae.*

The basis for most of our computations for PD(α,θ) with $0 < \alpha < 1$ is the following Proposition, according to which the PD(α,θ) distribution admits a density relative to the PD$(\alpha, 0)$ distribution that is just a constant times $L^{\theta/\alpha}$, where L is the local time variable introduced in Proposition 10.

PROPOSITION 14 [51]. *Let $0 < \alpha < 1$ and $\theta > -\alpha$. For every nonnegative product measurable function f,*

$$E_{\alpha,\theta}[f(V_1, V_2, \ldots)] = C_{\alpha,\theta} E_{\alpha,0}[L^{\theta/\alpha} f(V_1, V_2, \ldots)], \tag{42}$$

where $L := \lim_{n\to\infty} nV_n^{\alpha}$ as in (24) and

$$C_{\alpha,\theta} = \frac{1}{E_{\alpha,0}(L^{\theta/\alpha})} = \frac{\Gamma(\theta+1)}{\Gamma((\theta/\alpha)+1)}\Gamma(1-\alpha)^{\theta/\alpha}. \tag{43}$$

This proposition is a re-expression in terms of this paper of Corollary 3.15 of [51] (which contains misprints which should be corrected as follows: replace the first, third and fourth occurrences of B_p^{n*} by B_p^n). The constant $C_{\alpha,\theta}$ is determined by (30). See also [59, 52] for various alternative expressions for L.

Proposition 14 can be reformulated in various ways using different descriptions of PD$(\alpha, 0)$. For example, in the setting of Proposition 6, with $V_n(\tau_1)$ the nth largest jump of a stable (α) subordinator (τ_s) over $0 \le s \le 1$, we obtain

$$E_{\alpha,\theta}[f(V_1, V_2, \ldots)] = c_{\alpha,\theta} E\left[\tau_1^{-\theta} f\left(\frac{V_1(\tau_1)}{\tau_1}, \frac{V_2(\tau_1)}{\tau_1}, \ldots\right)\right], \tag{44}$$

where $c_{\alpha,\theta} = C^{\theta/\alpha} C_{\alpha,\theta}$ for $C_{\alpha,\theta}$ as in (43).

Proposition 14 shows that for fixed α with $0 < \alpha < 1$ the PD(α,θ) distributions are mutually absolutely continuous as θ varies. By contrast, for $\alpha = 0$ it is well known that the PD$(0,\theta)$ distributions are mutually singular as θ varies. Due to the way the definition of the local time variable L depends on α, the PD$(\alpha, 0)$ distributions are mutually singular as α varies, hence so too are the PD(α,θ) distributions for any fixed θ.

In Section 7 we obtain the following result, which generalizes both the Markov chain description of PD$(0,\theta)$ due to Vershik and Shmidt [66, 67] and Ignatov [34], and Proposition 8 for PD$(\alpha, 0)$. Note the parallel between (4) and (46).

THEOREM 15. *Let*

$$Y_n = V_n/(V_n + V_{n+1} + \cdots), \tag{45}$$

so

$$V_1 = Y_1, \qquad V_n = (1-Y_1)\cdots(1-Y_{n-1})Y_n \qquad (n \ge 2). \tag{46}$$

Let $R_n = V_{n+1}/V_n$. For $0 \le \alpha < 1$, $\theta > -\alpha$, let $P_{\alpha,\theta}$ govern (V_n) according to the PD(α,θ) *distribution and let $P^*_{\alpha,\theta}$ govern $(R_1, R_2, \ldots)$ as a sequence of independent random variables, such that R_n has* beta$(\theta+n\alpha, 1)$ *distribution. Then*

$$E_{\alpha,\theta}[f(Y_1, Y_2, \ldots)] = c_{\alpha,\theta} E^*_{\alpha,\theta}[Y_1^\theta f(Y_1, Y_2, \ldots)] \tag{47}$$

for a constant $K_{\alpha,\theta}$. Both $P_{\alpha,\theta}$ and $P^*_{\alpha,\theta}$ govern (Y_n) as a Markov chain with the same forward transition probabilities.

The chain (Y_n) is stationary and homogeneous under $P^*_{0,\theta}$, but for $0 < \alpha < 1$ the chain is nonhomogeneous, and the distribution of Y_n depends on n, in a manner described precisely in Section 7.

According to Proposition 8, under $P_{\alpha,0}$ for $0 < \alpha < 1$ the ratios $R_n := V_{n+1}/V_n$ are mutually independent. Under $P_{\alpha,\theta}$ for $\theta \ne 0$ this is no longer true. However, it follows from Theorem 15 that under $P_{\alpha,\theta}$ the R_n are asymptotically independent for large n with beta$(\theta+n\alpha, 1)$ distributions. There is also the following formula for the joint density of $R_1, \ldots, R_n$:

PROPOSITION 16. *Suppose $0 < \alpha < 1$, $\theta > -\alpha$ and $\theta \ne 0$. For $0 < r_i < 1$, $i = 1, 2, \ldots, n$,*

$$\frac{P_{\alpha,\theta}(R_1 \in dr_1, \ldots, R_n \in dr_n)}{dr_1 \cdots dr_n} = C_{\alpha,\theta}\, \alpha^n \Phi_\alpha\left(n + \frac{\theta}{\alpha}, \theta, a_n\right) \prod_{i=1}^n r_i^{i\alpha-1},$$

where

$$a_n = \frac{1}{r_n} + \frac{1}{r_n r_{n-1}} + \cdots + \frac{1}{r_n \cdots r_1}$$

and the function Φ_α is defined by

$$\begin{aligned} \Phi_\alpha(\ell, \xi, a) &:= \frac{\Gamma(\ell+1)}{\Gamma(\xi)} \int_0^\infty dt\, t^{\xi-1} e^{-t-at} \psi_\alpha(t)^{-\ell-1} \\ &= E_{\alpha,0}\left[\frac{L^\ell V_1^{\xi-\alpha\ell}}{(1+aV_1)^\xi}\right]. \end{aligned} \tag{48}$$

3.2. *One-dimensional distributions.* As an application of Proposition 14 we obtain the following formula for moments of the one-dimensional marginals of a PD(α,θ) distributed sequence:

PROPOSITION 17. *For $0 < \alpha < 1$, $\theta > -\alpha$, $p > 0$, $n = 1, 2 \ldots$,*

$$\begin{aligned} E_{\alpha,\theta}(V_n^p) &= \frac{\Gamma(1-\alpha)^{\theta/\alpha}}{\Gamma(n)} \frac{\Gamma(\theta+1)}{\Gamma(\theta+p)} \frac{\Gamma(\theta/\alpha+n)}{\Gamma(\theta/\alpha+1)} \\ &\quad \times \int_0^\infty dt\, t^{p+\theta-1} e^{-t} \phi_\alpha(t)^{n-1} \psi_\alpha(t)^{-\theta/\alpha-n}, \end{aligned} \tag{49}$$

where $\phi_\alpha(t)$ and $\psi_\alpha(t)$ are as in (33) *and* (34).

The following asymptotics as $n \to \infty$ are consequences of (24): for $0 < \alpha < 1$, $\theta > -\alpha,\ p > 0$,

$$n^{p/\alpha} E_{\alpha,\theta}(V_n^p) \to \frac{C_{\alpha,\theta}}{C_{\alpha,\theta+p}}, \tag{50}$$

where $C_{\alpha,\theta}$ is defined by (43) and the right-hand side of (50) is the pth moment of the $P_{\alpha,\theta}$ almost sure limit of $n^{1/\alpha} V_n$, that is, $L^{1/\alpha}$. Note from (42) that the $P_{\alpha,\theta}$ distribution of L has a strictly positive density $f_{\alpha,\theta}$ on $(0,\infty)$ given by $f_{\alpha,\theta}(\ell) = C_{\alpha,\theta}\ell^{\theta/\alpha} f_{\alpha,0}(\ell)$ where $f_{\alpha,0}(\ell)$ is the Mittag–Leffler density of the $P_{\alpha,0}$ distribution of $\Gamma(1-\alpha)L$, as discussed below (30).

By passage to the limit as $\alpha \downarrow 0$ (see Section 5.2), we recover from (49) the following known formula for PD$(0,\theta)$:

COROLLARY 18 [63, 73, 25, 50].

$$E_{0,\theta}(V_n^p) = \frac{\Gamma(\theta)}{\Gamma(\theta+p)} \frac{\theta^n}{\Gamma(n)} \int_0^\infty dt\, t^{p-1} e^{-t} E(t)^{n-1} e^{-\theta E(t)}, \tag{51}$$

where $E(t) = \int_t^\infty x^{-1} e^{-x}\, dx$.

The $P_{\alpha,\theta}$ distribution of V_n on $[0,1]$ is not easy to describe explicitly. There is, however, the following implicit description for $n = 1$:

PROPOSITION 19. *The $P_{\alpha,\theta}$ density of V_1 is uniquely determined for all $0 \le \alpha < 1$ and $\theta > -\alpha$ by the identity*

$$\begin{aligned} \frac{P_{\alpha,\theta}(V_1 \in dx)}{dx} &= \frac{\Gamma(\theta+1)}{\Gamma(\theta+\alpha)\Gamma(1-\alpha)} x^{-\alpha-1}(1-x)^{\alpha+\theta-1} \\ &\quad \times P_{\alpha,\alpha+\theta}\left(V_1 < \frac{x}{1-x}\right). \end{aligned} \tag{52}$$

The special case of (52) with $\alpha = 0$ and $\theta = 1$ appears as equation (3) of Vershik [65], attributed to Dickman [13]. See also [72, 26, 50] for alternative approaches to computation of the distribution of V_1 for PD$(0,\theta)$ and Lamperti [42] for a different functional equation that determines the distribution of $1/V_1$ for PD$(\alpha,0)$. In Section 8.1 a formula of Perman [50] is applied to obtain an expression for the $P_{\alpha,\theta}$ joint density of $V_1,\dots,V_n$ for $0 < \alpha < 1$, $\theta > -\alpha$, which is analogous to known results for PD$(0,\theta)$ [6, 66, 34]. In particular, this approach yields the following extension of results in Section 4 of [50] for the cases $\theta = 0$ and $\theta = \alpha$. To simplify notation, let $\bar{u} = 1 - u$.

PROPOSITION 20. *For all $0 \le \alpha < 1$ and $\theta > -\alpha$,*

$$\frac{P_{\alpha,\theta}(V_1 \in du)}{du} = \sum_1^\infty (-1)^{n+1} c_{n,\alpha,\theta} \frac{\bar{u}^{\alpha+\theta-1}}{u^{\alpha+1}} I_{n,\alpha,\theta}(u) \qquad (0 < u < 1), \tag{53}$$

where $I_{n,\alpha,\theta}(u) = 0$ *if* $u > 1/n$, *so all but the first* n *terms of the sum are zero if* $u > 1/(n+1)$, $I_{1,\alpha,\theta}(u) = 1$ *and, for* $n = 2, 3, \ldots$ *and* $0 < u_n \le 1/n$, $I_{n,\alpha,\theta}(u_n)$ *is the* $(n-1)$*-fold integral*

(54) $$I_{n,\alpha,\theta}(u_n) = \int \cdots \int \prod_{i=1}^{n-1} \frac{\bar{u}_i^{(n+1-i)\alpha+\theta-1}}{u_i^{\alpha+1}} 1\left(\frac{u_{i+1}}{\bar{u}_{i+1}} \le u_i \le \frac{1}{i}\right) du_i$$

and $c_{n,0,\theta} = \theta^n$ *while for* $0 < \alpha < 1$, $\theta > -\alpha$

(55) $$c_{n,\alpha,\theta} = \frac{\Gamma(\theta+1)\Gamma(\theta/\alpha+n)\alpha^{n-1}}{\Gamma(\theta+n\alpha)\Gamma(\theta/\alpha+1)\Gamma(1-\alpha)^n}.$$

For $1/2 < u < 1$ there is only one positive term in (53) and the formula reduces to (52). For $1/3 < u \le 1/2$ there are two nonzero terms in (53). This formula appears in the bridge case $\theta = \alpha$ at the bottom of page 278 of [50], but with a typographical error: $2\alpha\Gamma(\alpha)$ should be replaced by $2\alpha^2\Gamma(\alpha)$.

To illustrate using Proposition 7, for $\alpha = \theta = 1/2$, Proposition 20 describes the density of the length V_1 of the longest excursion of a Brownian bridge. Explicit integration is possible in this case at least for $n = 1, 2, 3$ to obtain

(56) $$P_{1/2,1/2}(V_1 \in du)/du = q_1(u) - q_2(u) + q_3(u) \quad \text{for } 1/4 < u < 1,$$

where the $q_n(u)$ are given for $0 < u < 1$ and $n = 1, 2, 3$ by

(57) $$q_1(u) = \tfrac{1}{2}u^{-3/2},$$

(58) $$q_2(u) = 1\left(u \le \frac{1}{2}\right)\frac{1}{\pi}u^{-3/2}\left(-\pi + 2\sqrt{\frac{1-2u}{u}} + 2\arcsin\sqrt{\frac{u}{1-u}}\right),$$

(59) $$q_3(u) = 1\left(u \le \frac{1}{3}\right)\frac{3}{4\pi}u^{-3/2}\left(2+2\pi+\frac{2}{u}-8\sqrt{\frac{1-2u}{u}}-8\arcsin\sqrt{\frac{u}{1-u}}\right).$$

See also Wendel [74] for another expression for the $P_{\alpha,\alpha}$ distribution of V_1 based on a method of Rosén, and see [39] for related results.

3.3. *A subordinator representation for* $0 < \alpha < 1$, $\theta > 0$. In view of Propositions 5 and 6, it is natural to look for a representation of PD(α, θ) as the distribution of the sequence

(60) $$\left(\frac{V_1(T)}{T}, \frac{V_2(T)}{T}, \ldots\right)$$

derived as in (7) from the ranked lengths $V_n(T)$ of component intervals of the set $[0, T]\backslash Z$, where Z is the closure of the range of a suitable subordinator $(\tau_s, s \ge 0)$ and T is a suitably defined random time. Such a representation is provided by the following Proposition. We write $\tau(s)$ instead of τ_s when typographically convenient.

PROPOSITION 21. *Fix α with $0 < \alpha < 1$ and $C > 0$. Let $(\tau_s, s \geq 0)$ be a subordinator with Lévy measure $\alpha C x^{-\alpha-1} e^{-x}\, dx$. Independent of $(\tau_s, s \geq 0)$, let $(\gamma(t), t \geq 0)$ be a gamma subordinator as defined below* (8). *For $\theta > 0$ let*

$$S_{\alpha,\theta} = \frac{\gamma(\theta/\alpha)}{C\Gamma(1-\alpha)}. \tag{61}$$

Then for $T = \tau(S_{\alpha,\theta})$ the sequence (60) *has* PD(α, θ) *distribution, independently of T, which has the same* gamma(θ) *distribution as $\gamma(\theta)$.*

Notice that in contrast to the formula of Proposition 14, all objects appearing in Proposition 21 have sensible limits as $\alpha \to 0$ for fixed θ. Take C so that $\alpha C \to 1$ as $\alpha \to 0$. Then as $\alpha \to 0$, the Lévy measure $\alpha C x^{-\alpha-1} e^{-x}\, dx$ of the subordinator (τ_s) approaches the Lévy measure $x^{-1} e^{-x}\, dx$ of a gamma process, while $S_{\alpha,\theta}$ converges in probability to the constant θ by the law of large numbers. So in the limit as $\alpha \to 0$ we recover Kingman's representation of PD$(0, \theta)$ stated in Proposition 5.

Proposition 21 is closely related to the following result, originally obtained by an entirely different argument. See also Proposition 33 below.

PROPOSITION 22 [53]. *For $0 < \alpha < 1$ and $\theta > 0$, suppose (U_n) has* PD$(0, \theta)$ *distribution, and independent of (U_n) let (V_{mn}), $m = 1, 2 \ldots,$ be a sequence of independent copies of (V_n) with* PD$(\alpha, 0)$ *distribution. Let (W_n) be defined by ranking the collection of products $\{U_m V_{mn}, m \in \mathbb{N}, n \in \mathbb{N}\}$. Then (W_n) has* PD(α, θ) *distribution.*

4. Development for PD($\boldsymbol{\alpha}$, 0).

4.1. *Proofs of the main results.* We will prove Proposition 10 first, then Proposition 8. Otherwise the proofs are in the same order as the propositions.

PROOF OF PROPOSITION 10. It is enough to establish the assertions (i), (ii) and (iii) of the proposition for any particular sequence (V_n) with PD$(\alpha, 0)$ distribution. We use $V_n := V_n(\tau_1)/\tau_1$ for a stable (α) subordinator (τ_s) as in (13). We first verify a modified form of the assertions (i),(ii) and (iii) in this case, with the definitions (25) replaced by

$$L := C\tau_1^{-\alpha}, \qquad \Sigma := \tau_1, \qquad \Delta_n := V_n(\tau_1). \tag{62}$$

The modified form of (ii) follows from the fact that the $V_n(\tau_1)$ are the ranked points of a PRM $\Lambda_\alpha(dx)$ on $(0, \infty)$. The modified form of (iii) follows by the usual change of variables to reduce the inhomogeneous PRM $\Lambda_\alpha(dy)$ on $(0, \infty)$ to a homogeneous PRM dx on $(0, \infty)$. Now (24) with a.s. convergence and $L = C\tau_1^{-\alpha}$ follows because $X_n/n \to 1$ a.s. by the law of large numbers. (This argument is due to Kingman [38]: our formula (24) is his (68)). See Section 4.3 for justification of the convergence (24) in pth mean. Tracing back through these definitions shows that the random variables defined in (62) can be recovered a.s. from L via (25). Thus (i),(ii) and (iii) hold for any (V_n) with PD$(\alpha, 0)$ distribution. □

PROOF OF PROPOSITION 8. By definition of R_n and the notation in Proposition 10,

$$R_n := \frac{V_{n+1}}{V_n} = \frac{\Delta_{n+1}}{\Delta_n} = \left(\frac{X_n}{X_{n+1}}\right)^{1/\alpha}. \tag{63}$$

Thus Proposition 8 reduces by a simple change of variables to the following elementary property of the points $0 < X_1 < X_2 < \cdots$ of a homogeneous Poisson process on $(0, \infty)$: the X_n/X_{n+1} are mutually independent beta$(n, 1)$ variables. □

We record for later use the following result, which is easily obtained by examination of the above proof:

COROLLARY 23. *In the setting of Proposition 10, for each $n = 1, 2, \ldots$ the random vector $(R_1, \ldots, R_n)$ is independent of the random sequence $(X_{n+1}, X_{n+2}, \ldots)$.*

The following lemma serves as a basis for further calculations.

LEMMA 24. *Let $\Delta_1 > \Delta_2 > \cdots$ be the ranked points of a* PRM $\Lambda_\alpha(dx)$ *on* $(0, \infty)$*, where $\Lambda_\alpha(x, \infty) = Cx^{-\alpha}$ for some $\alpha > 0$ and $C > 0$. Then:*

(i) *$C\Delta_n^{-\alpha}$ has* gamma(n) *distribution;*

(ii) *for $n \geq 2$ the $n - 1$ ratios*

$$\frac{\Delta_1}{\Delta_n} > \frac{\Delta_2}{\Delta_n} > \cdots > \frac{\Delta_{n-1}}{\Delta_n}$$

are distributed like the order statistics of $n - 1$ independent random variables with common distribution $C^{-1}\Lambda_\alpha(dx)1(x > 1)$, independently of $\Delta_n, \Delta_{n+1}, \ldots$;

(iii) *conditionally given $\Delta_1, \ldots, \Delta_n$ for $n \geq 1$, the*

$$\frac{\Delta_{n+1}}{\Delta_n} > \frac{\Delta_{n+2}}{\Delta_n} > \cdots$$

are the ranked points of a PRM $\Delta_n^{-\alpha}\Lambda_\alpha(dx)1(x < 1)$.

PROOF. Basic properties of Poisson processes imply that conditionally given $\Delta_n = a$, for $n \geq 2$ and $a > 0$, the $\Delta_1 > \Delta_2 > \cdots > \Delta_{n-1}$ are distributed like the order statistics of $n - 1$ independent random variables with common distribution $\Lambda_\alpha(a, \infty)^{-1}\Lambda_\alpha(dx)1(x > a)$; and conditionally given $\Delta_1, \ldots, \Delta_n$ for $n \geq 1$ with $\Delta_n = a$, the $\Delta_{n+1} > \Delta_{n+2} > \cdots$ are the ranked points of a PRM $\Lambda_\alpha(dx)1(x > a)$. Since under the transformation $u = x/a$ the image of the measure $\Lambda_\alpha(dx)$ is $a^{-\alpha}\Lambda_\alpha(du)$, the assertions of the lemma follow easily. □

PROOF OF PROPOSITION 11. Represent the PD$(\alpha, 0)$ distributed sequence (V_n) in terms of the points Δ_n of a PRM Λ_α as in Proposition 10. So

$$\frac{1}{V_n} = \frac{\Delta_1 + \cdots + \Delta_{n-1}}{\Delta_n} + \frac{\Delta_n}{\Delta_n} + \frac{\Delta_{n+1} + \Delta_{n+2} + \cdots}{\Delta_n} = A_{n-1} + 1 + \Sigma_n, \tag{64}$$

For $n \geq 2$ there is the representation

$$A_{n-1} = \frac{\Delta_1}{\Delta_n} + \frac{\Delta_2}{\Delta_n} + \cdots + \frac{\Delta_{n-1}}{\Delta_n}, \tag{65}$$

where the $(\Delta_i/\Delta_n, 1 \leq i \leq n-1)$ are distributed as the ranked values of $n-1$ independent random variables with the same distribution as A_1. Thus A_{n-1} is distributed like the sum of $n-1$ independent copies of $A_1 := \Delta_1/\Delta_2$, which has distribution

$$P(A_1 \in dx) = C^{-1}\Lambda_\alpha(dx)1(x > 1) = \alpha x^{-\alpha-1}dx1(x > 1). \tag{66}$$

This yields part (i). Consider now Σ_n defined by (32). Part (iii) of Lemma 24 represents Σ_n conditionally given $\Delta_1, \ldots, \Delta_n$ as the sum of points of a PRM $\Delta_n^{-\alpha}\Lambda_\alpha(dx)1(x < 1)$, whence

$$E[\exp(-\lambda\Sigma_n)|\Delta_1, \ldots, \Delta_n] = \exp\left(-\Delta_n^{-\alpha}\int_0^1 (1-\exp(-\lambda u))\Lambda_\alpha(du)\right). \tag{67}$$

Integration with respect to the gamma(n) distribution of $C\Delta_n^{-\alpha}$ yields (37), which establishes (ii). Finally, the independence claimed in part (iii) follows easily from the independence of the R_n. □

REMARK 25. The previous argument shows that for all $\alpha > 0$ formula (36) gives the Laplace transform of A_n defined by the last expression in (31) for a sequence of independent beta$(n\alpha, 1)$ distributed random variables (R_n), or by (65) in terms of Δ_n as in Lemma 24. However, the distribution of Σ_n is of interest only for $0 < \alpha < 1$, as it is easily seen that $\Sigma_n = \infty$ a.s. for $\alpha \geq 1$.

The following conditional form of Wendel's formula (38) proves useful in later calculations.

PROPOSITION 26. *Suppose (V_n) has* PD$(\alpha, 0)$ *distribution. Let (X_n), (R_n) and (A_n) be derived from (V_n) as in* (27), (21) *and* (31). *The conditional law of V_n given $R_1, \ldots, R_{n-1}$ and X_n is characterized by*

$$\begin{aligned} &E\left[\exp\left(-\frac{\lambda}{V_n}\right)\bigg| R_1, \ldots, R_{n-1}, X_n\right] \\ &\qquad = \exp(-\lambda(1 + A_{n-1}))\exp[-X_n(\psi_\alpha(\lambda) - 1)]. \end{aligned} \tag{68}$$

PROOF. Represent (V_n) in terms of the points (Δ_n) of a PRM Λ_α as in (26). Note that $\sigma(R_1, \ldots, R_{n-1}, X_n) = \sigma(\Delta_1, \ldots, \Delta_n)$ and use (35), (67) and $\Delta_n^{-\alpha} = X_n/C$. □

Consider now H_n derived as in (39) from the range of a stable (α) subordinator. Note that at time H_n the nth longest excursion interval that currently has length 1 is necessarily the meander interval. That is to say, $G_{H_n} = H_n - 1$, where for $t \geq 0$ we set

$$G_t = \sup(Z \cap [0, t)); \qquad D_t = \inf(Z \cap [t, \infty)). \tag{69}$$

Notice that H_n is just the nth instant t such that $t - G_t = 1$, so

$$0 < G_{H_1} < D_{H_1} < G_{H_2} < D_{H_2} < \cdots < G_{H_{n-1}} < D_{H_n} < G_{H_{n+1}}$$

and there is the natural decomposition

$$H_n = \sum_{j=1}^{n} (G_{H_j} - D_{H_{j-1}}) + \sum_{j=1}^{n-1} (D_{H_j} - G_{H_j}) + (H_n - G_{H_n}), \tag{70}$$

where $D_{H_0} = 0$ by convention, and the last term is $H_n - G_{H_n} = 1$. As shown by Wendel, formula (38) follows from the identity in distribution (40) and the observation that the first sum on the right-hand side of (70) is a sum of n independent terms with

$$G_{H_j} - D_{H_{j-1}} =_d G_{H_1} =_d \Sigma_1 \qquad (1 \le j \le n), \tag{71}$$

while the $n-1$ terms of the second sum in (70) are independent with

$$D_{H_j} - G_{H_j} =_d A_1 \qquad (1 \le j \le n-1), \tag{72}$$

where Σ_1 and A_1 are as in Proposition 11. These observations can be checked by repeated application of the strong Markov property at the times H_{D_j}, and the Poisson character of excursion interval lengths. Note that the $V_j(H_n)$ for $1 \le j \le n-1$ are the ranked values of the i.i.d. interval lengths $D_{H_j} - G_{H_j}$, $1 \le j \le n-1$, while $V_n(H_n) = 1$.

PROOF OF PROPOSITION 13. Let $(S_t, t \ge 0)$ denote the continuous local time process that is the inverse of the underlying stable (α) subordinator. The Poisson character of the interval lengths on the local time scale implies that for each fixed n the distribution of $[V_m(H_n), m = 1, 2, \ldots]$ can be described as follows:

(i) $V_n(H_n) = 1$;

(ii) for $0 < m < n$ the $V_m(H_n)$ are distributed like the order statistics of $m-1$ independent random variables with common distribution $C^{-1}\Lambda_\alpha(dx)1(x > 1)$;

(iii) independent of the $V_m(H_n)$ for $0 < m < n$, the multiple of the local time CS_{H_n} has a gamma(n) distribution;

(iv) given S_{H_n} and the $V_m(H_n)$ for $0 < m < n$, the $V_m(H_n)$ for $n < m < \infty$ are distributed as the ranked points of a PRM $S_{H_n}\Lambda_\alpha(dx)1(x < 1)$.

On the other hand, Lemma 24 shows that the same four statements hold if the following substitutions are made:

replace $V_m(H_n)$ by Δ_m/Δ_n and replace S_{H_n} by $\Delta_n^{-\alpha}$,

where the Δ_n are the ranked points of a PRM $\Lambda_\alpha(dx)$. Therefore, for each fixed $n = 1, 2, \ldots$,

$$\left(\frac{V_m(H_n)}{V_n(H_n)}, m = 1, 2, \ldots\right) =_d \left(\frac{\Delta_m}{\Delta_n}, m = 1, 2, \ldots\right). \tag{73}$$

The distribution of the sequence in (41) is now identified as PD$(\alpha, 0)$ using Proposition 10(ii). □

4.2. *A differential equation related to ϕ_α and ψ_α.* A proof of (36) can also be obtained using the recurrence relation

$$A_n = (1 + A_{n-1})/R_n, \tag{74}$$

together with the independence of A_{n-1} and R_n, the beta$(n\alpha, 1)$ distribution of R_n and the fact that

$$\exp(\lambda)\phi_\alpha(\lambda) = E\big[\exp\big(-\lambda\big(R_1^{-1} - 1\big)\big)\big] \tag{75}$$

solves the differential equation

$$\alpha = (\alpha + \lambda) f(\lambda) - \lambda f'(\lambda). \tag{76}$$

Another solution of (76) is the function

$$\exp(\lambda)\psi_\alpha(\lambda) = (E[\exp(-\lambda/V_1)])^{-1}. \tag{77}$$

In fact, all solutions of (76) are given by the formula

$$f(\lambda) = \lambda^\alpha e^\lambda \left[c + \alpha \int_\lambda^\infty \frac{dx\, e^{-x}}{x^{\alpha+1}} \right], \tag{78}$$

where $c = \lim_{\lambda\to\infty} \lambda^{-\alpha} e^{-\lambda} f(\lambda)$ is an arbitrary constant. Hence, $e^\lambda \phi_\alpha(\lambda)$ is the solution of (76) with $c = 0$, whereas $e^\lambda \psi_\alpha(\lambda)$ is the solution of (76) with $c = \Gamma(1 - \alpha)$, in agreement with formula (34). It can also be checked that the fact that $e^\lambda \psi_\alpha(\lambda)$ solves (76), together with the recurrence

$$\Sigma_n = R_n(1 + \Sigma_{n+1}), \tag{79}$$

where R_n and Σ_{n+1} are independent, is in agreement with formula (37). However, in contrast with the situation for (36), it seems difficult to prove (37) from this approach.

4.3. *Some absolute continuity relationships.* For $X_1 < X_2 < \cdots$ the points of a homogeneous Poisson process on $(0, \infty)$ with rate 1, there is the elementary absolute continuity relation

$$E[f(X_{m+1}, X_{m+2}, \ldots)] = \frac{1}{m!} E[X_1^m f(X_1, X_2, \ldots)], \tag{80}$$

where f is a generic positive measurable function of its arguments. For R_n as in (63), a change of variables yields

$$E[f(R_{m+1}, R_{m+2}, \ldots)] = \frac{1}{m!} E[X_1^m f(R_1, R_2, \ldots)], \tag{81}$$

where by a paraphrase of (24),

$$X_1 = \lim_{n\to\infty} n(R_1R_2\cdots R_n)^{\alpha} \quad \text{a.s.} \tag{82}$$

On the other hand, a direct calculation of the density ratio using Proposition 8 shows that

$$\begin{aligned} &E[f(R_{m+1}, R_{m+2}, \dots, R_{m+n})] \\ &\quad = \binom{n+m}{m} E[(R_1R_2\cdots R_n)^{m\alpha} f(R_1, R_2, \dots, R_n)]. \end{aligned} \tag{83}$$

Comparison of (81) and (83) shows that

$$E\left(\frac{X_1^m}{m!}\middle| R_1, R_2, \dots, R_n\right) = \binom{n+m}{m}(R_1R_2\cdots R_n)^{m\alpha}. \tag{84}$$

Since $X_1/(R_1R_2\cdots R_n)^{\alpha} = X_{n+1}$ this amounts to the consequence of Corollary 23 and (63) that X_{n+1} is independent of $(R_1, R_2, \dots, R_n)$ and is distributed as gamma$(n+1)$. Since X_1 has finite moments of all orders, martingale convergence shows that the a.s. convergence in (82) takes place also in pth mean for every $p \geq 1$. It follows easily that the same is true of the a.s. convergence in (24).

5. Development for PD($\boldsymbol{\alpha}$, $\boldsymbol{\theta}$).

5.1. *Proofs of some results.*

PROOF OF PROPOSITION 17. Combine Proposition 14 and the following lemma. □

LEMMA 27. *Suppose* (V_n) *has* PD$(\alpha, 0)$ *distribution and let* $L = \lim_n nV_n^{\alpha}$ *as in* (24)*. Then for all real* $\ell > -1$ *and* $p > 0$*, and* $n = 1, 2\dots,$

$$E[L^{\ell} V_n^p] = \frac{\Gamma(\ell+n)}{\Gamma(n)\Gamma(p+\ell\alpha)} \int_0^{\infty} dt\, t^{p+\ell\alpha-1} e^{-t} \phi_{\alpha}(t)^{n-1} \psi_{\alpha}(t)^{-\ell-n}. \tag{85}$$

PROOF. We will use the following expression for negative moments of a positive random variable X in terms of its Laplace transform:

$$E[X^{-p}] = \frac{1}{\Gamma(p)} \int_0^{\infty} dt\, t^{p-1} E[e^{-tX}] \qquad (p > 0). \tag{86}$$

Combined with Wendel's formula (38), this immediately yields the special case of (85) with $\ell = 0$. Recall that $X_n := LV_n^{-\alpha}$, so the left-hand side of (85) is

$$E(X_n^{\ell} V_n^{p+\ell\alpha}) = E\left[\frac{X_n^{\ell}}{\Gamma(p+\ell\alpha)} \int_0^{\infty} dt\, t^{p+\ell\alpha-1} E\left[\exp\left(-\frac{t}{V_n}\right)\middle| X_n\right]\right]$$

by (86). Now use (68) and the fact that X_n has gamma(n) distribution independent of A_{n-1} to obtain by elementary integration

$$E[L^\ell V_n^p] = \frac{1}{\Gamma(n)\Gamma(p+\ell\alpha)}\int_0^\infty dt\, t^{p+\ell\alpha-1}\exp(-t)\Gamma(\ell+n)\psi_\alpha(t)^{-\ell-n}E[\exp(-tA_{n-1})].$$

Here, for $n=1$, $A_0=0$. Now use (36) to obtain (85). □

REMARK 28. Consider (85) for $p=0, \ell>0$. Since the left-hand side does not depend on n, neither does the right, something which is not evident a priori. This can be shown to be equivalent to the Wronskian identity

$$(\phi_\alpha\psi'_\alpha - \psi_\alpha\phi'_\alpha)(t) = \alpha\Gamma(1-\alpha)e^{-t}t^{\alpha-1}, \tag{87}$$

which follows from the description of ϕ_α and ψ_α in terms of the differential equation (76).

Further moment formulae. Suppose (V_n) has PD$(\alpha,0)$ distribution. Let L, X_n, R_n, A_n and Σ_n be the random variables defined in terms of (V_n) as in (24), (27), (21), (31) and (32).

As a first variant of (85), we can compute similarly

$$E\left[L^\ell V_n^p\exp\left(-\frac{\lambda}{V_n}\right)\right] = E\left[X_n^\ell V_n^{p+\ell\alpha}\exp\left(-\frac{\lambda}{V_n}\right)\right] = E\left[\frac{X_n^\ell}{\Gamma(p+\ell\alpha)}\int_0^\infty dt\, t^{p+\ell\alpha-1}E\left[\exp\left(-\frac{(t+\lambda)}{V_n}\right)\middle| X_n\right]\right].$$

Using (68) and then (36) again, with $t+\lambda$ instead of λ, yields

$$E\left[L^\ell V_n^p\exp\left(-\frac{\lambda}{V_n}\right)\right] = \frac{\Gamma(\ell+n)}{\Gamma(n)\Gamma(p+\ell\alpha)}\int_0^\infty dt\, t^{p+\ell\alpha-1}\exp(-t-\lambda)\phi_\alpha(t+\lambda)^{n-1}\psi_\alpha(t+\lambda)^{-\ell-n}. \tag{88}$$

PROOF OF PROPOSITION 16. This follows easily from Propositions 8 and 14 using the following lemma, which states another variant of (85):

LEMMA 29. *Suppose* (V_n) *has* PD$(\alpha,0)$ *distribution. Let* $L := \lim_{n\to\infty} nV_n^\alpha$ *as in* (24) *and let* $R_n := V_{n+1}/V_n$. *For all real* $\ell > -1$ *and* $\gamma > 0$, *and* $n = 0,1,2,\dots,$

$$E[L^\ell V_1^{\gamma-\alpha\ell} \mid R_1,\dots,R_n] = \frac{1}{n!}\left(\prod_{j=1}^n R_j\right)^{\ell\alpha-\gamma}\Phi_\alpha(\ell+n,\gamma,A_n) \tag{89}$$

for A_n *as in* (31) *and* $\Phi_\alpha(\ell,\gamma,a)$ *as in* (48).

PROOF. Let $\mathscr{R}_n = \sigma(R_1, \dots, R_n)$. Elementary manipulations show that

$$E[L^\ell V_1^{\gamma-\alpha\ell} \,|\mathscr{R}_n] = \left(\prod_{j=1}^n R_j\right)^{\ell\alpha-\gamma} E[(L/V_{n+1}^\alpha)^\ell\, V_{n+1}^\gamma|\mathscr{R}_n]. \tag{90}$$

Now use (86) for $p = \gamma$ and $X = 1/V_{n+1}$ to express the right-hand side of (90) as

$$\left(\prod_{j=1}^n R_j\right)^{\ell\alpha-\gamma} \frac{1}{\Gamma(\gamma)} \int_0^\infty dt\, t^{\gamma-1}\{\cdots\},$$

where

$$\{\cdots\} = E[(L/V_{n+1}^\alpha)^\ell\, E[\exp(-t/V_{n+1})|\mathscr{R}_n, X_{n+1}] \mid \mathscr{R}_n]$$

and $X_{n+1} := LV_{n+1}^{-\alpha}$. Now use formula (68) to show that

$$\begin{aligned}\{\cdots\} &= E[X_{n+1}^\ell \exp(X_{n+1}(1-\psi_\alpha(t)))]\exp(-t(1+A_n))\\ &= \exp(-t(1+A_n))\,\frac{\Gamma(n+\ell)}{n!}\,\psi_\alpha(t)^{-(\ell+n+1)}\end{aligned}$$

by the independence of A_n and X_{n+1} [see Corollary 23 and (31)] and elementary integration with respect to the gamma$(n+1)$ distribution of X_{n+1}. This yields formula (89) with Φ_α defined by (48). The second equality in (48) is easily obtained by another manipulation like (86). □

REMARK 30. It is also possible to derive (89), with Φ_α defined by the second expression in (48), by starting from Perman's formula for the joint density of Σ, $V_1, \dots, V_{n+1}$ stated in Proposition 47, and making suitable changes of variables and integrating out Σ and V_1.

PROOF OF PROPOSITION 19. For $(\tilde{V}_n)$ the size-biased permutation of (V_n) as in Definition 1 and Proposition 2, we can compute $P_{\alpha,\theta}(V_1 \in dx, V_1 = \tilde{V}_1)$ in two different ways. First, by conditioning on V_1 and using (2):

$$P_{\alpha,\theta}(V_1 \in dx, V_1 = \tilde{V}_1) = xP_{\alpha,\theta}(V_1 \in dx). \tag{91}$$

However, conditioning instead on $\tilde{V}_1$ and using the consequence of (4) that the $P_{\alpha,\theta}$ distribution of $(\tilde{V}_2, \tilde{V}_3, \dots)/(1-\tilde{V}_1)$ is identical to the $P_{\alpha,\alpha+\theta}$ distribution of $(\tilde{V}_1, \tilde{V}_2, \dots)$ yields

$$\begin{aligned}&P_{\alpha,\theta}(V_1 \in dx, V_1 = \tilde{V}_1)\\ &\quad= P_{\alpha,\theta}(\tilde{V}_1 \in dx, \max_{n\ge 2} \tilde{V}_n < x)\\ &\quad= P_{\alpha,\theta}(\tilde{V}_1 \in dx)P_{\alpha,\theta}\left(\max_{n\ge 2} \frac{\tilde{V}_n}{1-\tilde{V}_1} < \frac{x}{1-x}\middle|\tilde{V}_1 = x\right)\\ &\quad= \frac{\Gamma(\theta+1)}{\Gamma(\theta+\alpha)\Gamma(1-\alpha)} x^{-\alpha}(1-x)^{\alpha+\theta-1}\,dx\,P_{\alpha,\alpha+\theta}\left(V_1 < \frac{x}{(1-x)}\right).\end{aligned} \tag{92}$$

Now comparison of (91) and (92) yields (52). For $1/2 < x < 1$ it is obvious that $P_{\alpha,\alpha+\theta}(V_1 < x/(1-x)) = 1$, so (52) determines the $P_{\alpha,\theta}$ density of V_1 at x for $1/2 < x < 1$. [This case of (52) can also be read from (6)]. Recursive application of (52) now determines the $P_{\alpha,\theta}$ density of V_1 at x for $1/(n+1) < x < 1/n$, $n = 2, 3, \ldots$. □

PROOF OF PROPOSITION 21. Let $K = C\Gamma(1-\alpha)$. Let (σ_s) be a stable (α) subordinator with $E[\exp(-\lambda\sigma_s)] = \exp(-K\lambda^\alpha s)$. The Lévy measure of (τ_s) has density e^{-x} relative to the Lévy measure of (σ_s), which implies (see, e.g., [40]) that for each $s > 0$ and every positive measurable functional F,

$$E[\,F(\tau_t, 0 \le t \le s)\,] = \exp(Ks)E[\,F(\sigma_t, 0 \le t \le s)\exp(-\sigma_s)]. \tag{93}$$

Let $(V_1, V_2, \ldots)$ denote a sequence with PD(α, θ) distribution. Let L be the local time variable derived from $(V_1, V_2, \ldots)$ as in (24) and $\Sigma = (C/L)^{1/\alpha}$. From Propositions 14 and 10, the conditional law of $(V_1, V_2, \ldots)$ given $\Sigma = t$ does not depend on θ. Call it PD$(\alpha|\,t)$, say:

$$\text{PD}(\alpha|\,t) = \text{the conditional law of } \left(\frac{\Delta_1}{\sigma_1}, \frac{\Delta_2}{\sigma_1}, \ldots\right) \text{ given } \sigma_1 = t, \tag{94}$$

where $\Delta_1 > \Delta_2 > \cdots$ are the ranked jumps of $(\sigma_s, 0 \le s \le 1)$. Then from (44),

$$\text{PD}(\alpha, \theta) = c_{\alpha,\theta} \int_0^\infty \text{PD}(\alpha|\,t)t^{-\theta}P(\sigma_1 \in dt). \tag{95}$$

The finite-dimensional distributions of PD$(\alpha|\,t)$ are described by Perman's formula (153), but this description is not required in the following argument.

Let

$$\mathrm{W}_s = \left(\frac{V_1(\tau_s)}{\tau_s}, \frac{V_2(\tau_s)}{\tau_s}, \ldots\right).$$

From (93) and scaling properties of (σ_s) we learn that if ζ is a positive random variable independent of $(\tau_s, s \ge 0)$, then

$$\text{the conditional law of } \mathrm{W}_\zeta \text{ given } \zeta \text{ and } \tau_\zeta \text{ is PD}(\alpha|\,\tau_\zeta/\zeta^{1/\alpha}) \tag{96}$$

no matter what the distribution of ζ. Consequently

$$\zeta \text{ and } \mathrm{W}_\zeta \text{ are conditionally independent given } \tau_\zeta/\zeta^{1/\alpha}. \tag{97}$$

From (95) and (97), it now suffices to show that for $\zeta = K^{-1}\gamma(\theta/\alpha)$ the following three things are true:

$$P[\tau_\zeta/\zeta^{1/\alpha} \in dt\,] = c_{\alpha,\theta}\,t^{-\theta}P(\sigma_1 \in dt), \tag{98}$$

$$\tau_\zeta \text{ has gamma}(\theta) \text{ distribution}, \tag{99}$$

$$\tau_\zeta/\zeta^{1/\alpha} \text{ and } \tau_\zeta \text{ are independent}. \tag{100}$$

However, (98), (99) and (100) follow at once from the next lemma applied with $h(z) = cz^b$ for $b = \theta/\alpha$ and a constant c.

LEMMA 31. *Let $(\tau_s, s \geq 0)$ be as in Proposition 21 and let ζ be a random variable independent of $(\tau_s, s \geq 0)$ with density of the form*

$$P(\zeta \in dz) = h(z) \exp(-Kz) \frac{dz}{z} \tag{101}$$

for some function $h(z)$. Then for $t > 0$, $u > 0$,

$$P\left(\tau_\zeta \in du,\ \frac{\tau_\zeta}{\zeta^{1/\alpha}} \in dt\right) = \alpha e^{-u} h\left(\left(\frac{u}{t}\right)^\alpha\right) \frac{du}{u} P(\sigma_1 \in dt). \tag{102}$$

PROOF. Conditioning on $\zeta = z$, there is the following identity for all positive measurable functions f and g:

$$\begin{aligned} E\left[f(\tau_\zeta) g\left(\frac{\tau_\zeta}{\zeta^{1/\alpha}}\right) \middle| \zeta = z \right] &= E[f(\tau_z) g(\tau_z / z^{1/\alpha})] \\ &= \exp(Kz) E[f(z^{1/\alpha}\sigma_1) g(\sigma_1) \exp(-z^{1/\alpha}\sigma_1)] \end{aligned}$$

by (93) and the scaling property of the stable subordinator $(\sigma_s, s \geq 0)$. Integrate with respect to the distribution (101) of ζ to obtain

$$\begin{aligned} E\left[f(\tau_\zeta) g\left(\frac{\tau_\zeta}{\zeta^{1/\alpha}}\right)\right] &= \int_0^\infty \frac{dz}{z} h(z) E[f(z^{1/\alpha}\sigma_1) g(\sigma_1) \exp(-z^{1/\alpha}\sigma_1)] \\ &= E\left[\int_0^\infty \frac{du}{u} \alpha \exp(-u) h\left(\left(\frac{u}{\sigma_1}\right)^\alpha\right) f(u) g(\sigma_1)\right] \end{aligned} \tag{103}$$

by Fubini's theorem and the change of variable

$$u = z^{1/\alpha}\sigma_1, \qquad z = (u/\sigma_1)^\alpha, \qquad \frac{dz}{z} = \frac{du}{u}.$$

Now (103) amounts to (102). □

REMARK 32. Conversely, formula (102) shows that if any of (98), (99) or (100) holds, the function $h(z)$ introduced in (101) must be of the form $h(z) = cz^b$, that is, $K\zeta$ must have gamma(b) distribution for some $b > 0$. Consider for instance (100). From (102), for (100) to be satisfied, it is necessary that

$$h(u/v) = j(u)k(v) \quad \text{a.e. with respect to } du\,dv$$

for some functions j and k, hence that

$$h(uw) = c\,h(u)h(w) \quad \text{a.e. with respect to } du\,dw,$$

which forces $h(u) = cu^b$ for some c and b.

PROOF OF PROPOSITION 22. Proposition 22 follows from Proposition 21 and the next proposition, which in fact allows either of Propositions 22 or 21 to be derived easily from the other.

PROPOSITION 33. *In the setting of Proposition 21, let* $\zeta_t = K^{-1}\gamma(t)$, *where* $K = C\Gamma(1-\alpha)$, *and let* $S_1 > S_2 > \cdots$ *denote the ranked values of the jumps of* $(\zeta_t, 0 \le t \le \theta/\alpha)$, *say* $S_i = \zeta_{\tau_i} - \zeta_{\tau_i-}$, *where* τ_i *is the time of the jump of magnitude* S_i. *Let* $T_i = \tau(\zeta_{\tau_i}) - \tau(\zeta_{\tau_i-})$ *Then:*

(i) *the* (S_i, T_i), $i = 1, 2\dots$, *are the points of a* PRM *with intensity measure*

$$M(ds, dt) = \frac{\theta}{\alpha}\frac{ds}{s} f_s(t) e^{-t}\, dt = \theta \frac{dt}{t} e^{-t} g_t(s)\, ds, \tag{104}$$

where $f_s(t) = P(\sigma_s \in dt)/dt$ *and* $g_t(s) = P(S_t \in ds)/ds$, *where* $(S_t, t \ge 0)$ *is the inverse of the stable* (α) *subordinator* $(\sigma_s, s \ge 0)$.

(ii) *Let* $T_{\pi(i)}$ *be the* i*th largest of the jumps* T_i, $i = 1, 2, \dots$. *Then*

$$\left(\frac{T_{\pi(i)}}{\tau(\zeta_{\theta/\alpha})}, i = 1, 2, \dots\right) \text{has PD}(0, \theta) \text{ distribution}$$

independently of the gamma(θ) *variable* $\sum_i T_i = \tau(\zeta_{\theta/\alpha})$.

(iii) *if* $\Delta_{i1} > \Delta_{i2} > \cdots$ *are the ranked jumps of* (τ_s) *incurred over the* s-*interval whose length is* $S_{\pi(i)}$, *then for each* i *the sequence*

$$\left(\frac{\Delta_{ij}}{T_{\pi(i)}}, j = 1, 2, \dots\right) \text{ has PD}(\alpha, 0) \text{ distribution.}$$

Moreover these sequences are mutually independent as i *varies and independent also of the sequence* $(T_{\pi(i)}, i = 1, 2, \dots)$, *where*

$$T_{\pi(i)} = \Delta_{i1} + \Delta_{i2} + \cdots \quad \text{and} \quad \tau(\zeta_{\theta/\alpha}) = \sum_i T_{\pi(i)} = \sum_i \sum_j \Delta_{ij}$$

and the $V_n(\zeta_{\theta/\alpha})$ *featured in Proposition 21 are the ranked values of the* Δ_{ij}.

PROOF. Due to the Poisson character of the jumps of the two independent subordinators, the points (S_i, T_i), $i = 1, 2\dots$, are the points of a PRM with intensity measure

$$M(ds, dt) = \frac{\theta}{\alpha}\frac{ds}{s} \exp(-Ks) P(\tau_s \in dt), \tag{105}$$

which can be expressed as in (104) using (93) and the formula $f_s(t) = \alpha s g_t(s)/t$, which is a consequence of the identity in distribution $S_t/t^\alpha =_d s/\sigma_s^\alpha$ (see, e.g., Section 7 of [59]). This yields (i). Since $\int_0^\infty g_t(s)\, ds = 1$, the T_i are the points of a PRM $\theta t^{-1} e^{-t}\, dt$ over $t > 0$. So (ii) follows from Proposition 5. Turning to (iii), the last expression for $M(ds, dt)$ in (105), combined with standard facts about Poisson processes, shows that conditionally given all the $T_{\pi(i)}$, the corresponding jumps $S_{\pi(i)}$ of the gamma process $(\zeta_t, 0 \le t \le \theta/\alpha)$ are mutually independent, with

$$P(S_{\pi(i)} \in ds \mid T_{\pi(i)} = t) = g_t(s)\, ds.$$

Now (iii) follows using (96) and (95) for $\theta = 0$. □

5.2. *Limits as* $\alpha \to 0$. Let $\mathscr{P}$ denote the space of probability measures on $[0,1] \times [0,1] \times \cdots$ and give $\mathscr{P}$ the topology of weak convergence of finite dimensional distributions. It is immediate from Definition 1 that the $P_{\alpha,\theta}$ distribution of $(\tilde{V}_n)$ defines a continuous map from $\{(\alpha, \theta): 0 \le \alpha < 1, \theta > -\alpha\}$ to $\mathscr{P}$. As a consequence [16], the same is true of the $P_{\alpha,\theta}$ distribution of (V_n). That is to say, PD(α, θ) is continuous in (α, θ). In particular, for each $\theta > 0$ the limit of PD(α, θ) as $\alpha \downarrow 0$ is PD$(0, \theta)$. That is, for every bounded continuous function f defined on $[0,1]^n$,

$$\lim_{\alpha \downarrow 0} E_{\alpha,\theta}[f(V_1, \dots, V_n)] = E_{0,\theta}[f(V_1, \dots, V_n)]. \tag{106}$$

Proposition 21 provides a setting in which (106) follows from weak convergence as $\alpha \downarrow 0$ of a subordinator with Lévy measure $x^{-\alpha-1}e^{-x}\,dx$ to a gamma process with Lévy measure $x^{-1}e^{-x}\,dx$. See [68] for further discussion and [9] for other aspects of the asymptotic behavior of a stable (α) subordinator as $\alpha \downarrow 0$.

To illustrate (106), we now derive the known formula for $E_{0,\theta}(V_n^p)$ for $p > 0$ given in Corollary 18 from the corresponding formula for $E_{\alpha,\theta}(V_n^p)$ with $0 < \alpha < 1$ stated in Proposition 17.

Derivation of Corollary 18 *from Proposition* 17. The evaluation of the limit is justified by the following asymptotics as $\alpha \downarrow 0$:

$$\Gamma(1-\alpha)^{\theta/\alpha} \sim (1+\gamma\alpha)^{\theta/\alpha} \to e^{\theta\gamma}, \tag{107}$$

where $a(\alpha) \sim b(\alpha)$ means $a(\alpha)/b(\alpha) \to 1$ as $\alpha \downarrow 0$,

$$\gamma = -\Gamma'(1) \tag{108}$$

is Euler's constant and

$$\frac{\Gamma(\theta/\alpha + n)}{\Gamma(\theta/\alpha + 1)} \sim \frac{\theta^{n-1}}{\alpha^{n-1}}. \tag{109}$$

The factor α^{n-1} in the denominator is asymptotically cancelled inside the integral by the factor

$$\phi_\alpha(t)^{n-1} = \left(\alpha t^\alpha \int_t^\infty dx\, x^{-\alpha-1}e^{-x}\right)^{n-1} \sim \alpha^{n-1}E(t)^{n-1}. \tag{110}$$

Finally, in view of (108) and (110) for $n = 2$, formula (34) implies

$$\psi_\alpha(t) - 1 \sim \alpha(E(t) + \gamma + \log(t)) \tag{111}$$

and consequently

$$\begin{aligned} \psi_\alpha(t)^{-n-(\theta/\alpha)} &\to \exp(-\theta(E(t) + \gamma + \log(t))) \\ &= t^{-\theta}\exp(-\gamma\theta)\exp(-\theta E(t)). \end{aligned} \tag{112}$$

It is easily argued that these limiting operations can be switched with the integral in (49), and (51) results after some cancellation. □

6. Sampling from PD($\boldsymbol{\alpha}, \boldsymbol{\theta}$). Applications of a random discrete distribution (V_n) often involve a *sample from* (V_n), that is, a random variable N such that the conditional distribution of N given (V_n) is given by

(113) $$P(N = n | V_1, V_2, \ldots) = V_n \qquad (n = 1, 2, \ldots).$$

Then V_N is a size-biased pick from (V_n), as in (2). See for instance [73, 25] for a nice interpretation of N in the application of PD$(0, \theta)$ to population genetics.

6.1. *Deletion and insertion operations.* Given a sequence (v_n) and an index N, say (v'_n) is derived from (v_n) by *deletion of* v_N if

$$v'_n = v_n 1(n < N) + v_{n+1} 1(n \geq N).$$

The next proposition follows immediately from Proposition 2:

PROPOSITION 34. *Let N be a sample from (V_n) with* PD(α, θ) *distribution, where $0 \leq \alpha < 1$ and $\theta > -\alpha$. Let (V'_n) be derived from (V_n) by deletion of V_N, and let $V''_n = V'_n/(1 - V_N)$, $n = 1, 2, \ldots$. Then (V''_n) has* PD$(\alpha, \theta + \alpha)$ *distribution, independently of V_N, which has* beta$(1 - \alpha, \theta + \alpha)$ *distribution.*

In particular the PD$(0, \theta)$ distribution is invariant under this operation of size-biased deletion and renormalization, a result which is a known characterization of PD$(0, \theta)$ [48, 29].

Suppose a PD$(\alpha, 0)$ distributed sequence (V_n) has been constructed by any of the methods described in Section 2. By the operation of size-biased deletion and renormalization as above, we obtain a sequence with PD(α, α) distribution. Repeating the operation yields sequences with distributions PD$(\alpha, 2\alpha)$, PD$(\alpha, 3\alpha), \ldots$.

This result about deletion can be rephrased as a result about insertion: given $(v'_1 \geq v'_2 \geq \cdots)$ and a real number $v > \inf_n v'_n$, say (v_n) is derived from (v'_n) by *insertion of* v if

$$v_n = v'_n 1(n < N) + v 1(n = N) + v'_{n+1} 1(n > N),$$

where $N - 1 = \sum_{n=1}^{\infty} 1(v'_n > v)$ is the number of terms of (v'_n) that strictly exceed v. Note that $v_N = v$ by definition.

PROPOSITION 35. *Fix $0 \leq \alpha < 1$ and $\theta > -\alpha$. Let (V''_n) have* PD$(\alpha, \alpha + \theta)$ *distribution. Independent of (V''_n) let X have* beta$(1 - \alpha, \theta + \alpha)$ *distribution. Let (V_n) be defined by insertion of X into $((1 - X)V''_n, n = 1, 2, \ldots)$. Then (V_n) has* PD(α, θ) *distribution and $X = V_N$, where N is a sample from (V_n).*

6.2. *Distribution of a sample from* PD(α, θ). Immediately from (113), the unconditional distribution of a sample N from (V_n) with PD(α, θ) distribution is given by $P_{\alpha,\theta}(N = n) = E_{\alpha,\theta}(V_n)$ as specified in formulae (49) and (51) for $p = 1$. For PD($0, \theta$) this result is due to Griffiths [25]. Inspection of formula (51) for $p = 1$ shows that Griffiths' result can be restated as follows:

> for N a sample from PD($0, \theta$), the distribution of $N - 1$ is a mixture of Poisson (μ) distributions, with the parameter μ distributed as $\theta\Lambda(T, \infty)$, where $\Lambda(dx) = x^{-1}e^{-x}dx$ is the Lévy measure of a gamma subordinator, and T is a standard exponential variable.

This result can be understood probabilistically as follows, by application of Propositions 5 and 35. Take (V_n'') in Proposition 35 to be the PD($0, \theta$) sequence $V_n'' = V_n(\tau_\theta)/\tau_\theta$ derived from a gamma subordinator $(\tau_s, 0 \le s \le \theta)$ as in (10). Let $X = T/(T + \tau_\theta)$ for T a standard exponential independent of (τ_s) and let (V_n) be constructed as in Proposition 35. Let N be the rank of X in (V_n). According to Proposition 35, N is a sample from the PD($0, \theta$) sequence (V_n). However, by construction, $N - 1$ is the number of n such that $V_n(\tau_\theta) > T$, and given T this number has Poisson distribution with mean $\theta\Lambda(T, \infty)$.

The analog for $0 < \alpha < 1$ of the above result for PD($0, \theta$) is the subject of the next proposition.

PROPOSITION 36. *For each* $0 < \alpha < 1, \theta > -\alpha$, *the* $P_{\alpha,\theta}$ *distribution of* $N-1$ *is an integral mixture of negative binomial distributions with parameters* $\theta/\alpha + 1$ *and* p, *with a mixing distribution over* p *which depends only on* α. *More precisely, for each* $m = 0, 1, \dots,$

$$
\begin{aligned}
P_{\alpha,\theta}(N - 1 = m) &= E_{\alpha,\theta}(V_{m+1}) \\
&= \int_0^\infty P(Z_{1-\alpha} \in dz)\binom{\theta/\alpha + m}{m}(1 - p_\alpha(z))^m p_\alpha(z)^{\theta/\alpha+1},
\end{aligned}
\tag{114}
$$

where $Z_{1-\alpha}$ *has* gamma($1 - \alpha$) *distribution and*

$$
p_\alpha(z) = \frac{\psi_\alpha(z) - \phi_\alpha(z)}{\psi_\alpha(z)} = \frac{\Gamma(1-\alpha)z^\alpha}{\psi_\alpha(z)}
\tag{115}
$$

is such that $0 < p_\alpha(z) < 1$ *for all* $0 < \alpha < 1$ *and* $z > 0$.

PROOF. This can be obtained either by manipulation of formula (49) for $p = 1$, or more probabilistically by application of Proposition 35, as in the case $\alpha = 0$ discussed above, using the construction of Proposition 21 instead of Proposition 5. □

From (114) and the formula $r(1-p)/p$ for the mean of the negative binomial (r, p) distribution, for $0 < \alpha < 1$, $\theta > -\alpha$, there is the following formula for the mean of N:

$$
E_{\alpha,\theta}(N) = 1 + \left(1 + \frac{\theta}{\alpha}\right)\Gamma(1-\alpha)^{-2}\int_0^\infty dz\, z^{-2\alpha}e^{-z}\phi_\alpha(z),
\tag{116}
$$

which is linear in θ for fixed $\alpha < 1/2$ and infinite for all $\theta > -\alpha$ if $\alpha \geq 1/2$. Formulae for higher moments of N follow similarly, while asymptotics for $P_{\alpha,\theta}(N = n)$ and $P_{\alpha,\theta}(N \geq n)$ for large n are immediate from (50).

The next two sections illustrate two interesting special cases of Proposition 36 with natural interpretations in terms of excursions of a Brownian motion or Bessel process. We thank Yuval Peres and Steve Evans for a conversation which helped us develop these interpretations.

6.3. *The rank of the excursion in progress.* Consider the setup of Section 1.2, with Z the range of a stable (α) subordinator, and $V_n(t)$ the length of the nth longest interval component of $[0, t]\backslash Z$. So Z could be the zero set of Brownian motion $(\alpha = 1/2)$ or a recurrent Bessel process of dimension $2 - 2\alpha$ for $0 < \alpha < 1$. Let N_t be the rank of the meander length $t - G_t$ in the sequence of excursion lengths $V_1(t) > V_2(t) > \cdots$, so $t - G_t = V_{N_t}(t)$. According to Theorem 1.2 of [59], for each fixed time t the random variable N_t is a sample from $(V_n(t)/t)$. Combined with (14), this shows that the joint law of N_t and the sequence $(V_n(t)/t)$ is given by the formula

$$E\left[1(N_t = n) f\left(\frac{V_1(t)}{t}, \frac{V_2(t)}{t}, \ldots\right)\right] = E_{\alpha,0}[V_n f(V_1, V_2, \ldots)] \tag{117}$$

for all $n = 1, 2 \ldots$ and all nonnegative product measurable functions f. Here E denotes expectation relative to P governing the stable (α) subordinator (τ_s), and $E_{\alpha,0}$ denotes expectation relative to $P_{\alpha,0}$ governing (V_n) with PD$(\alpha, 0)$ distribution. In particular, from Proposition 36 for $0 < \alpha < 1$ and $\theta = 0$ we obtain for all $t > 0$,

$$P(N_t = n) = \int_0^\infty dz\, e^{-z} \phi_\alpha(z)^{n-1} \psi_\alpha(z)^{-n}. \tag{118}$$

This is a companion of a result of Scheffer [62], which can be expressed in present notation as

$$P(N_{D_t} = n) = \alpha \int_0^\infty dz\, z^{-1}(1 - e^{-z}) \phi_\alpha(z)^{n-1} \psi_\alpha(z)^{-n}. \tag{119}$$

Here, $N_t - 1$ is the number of excursions completed by time t whose lengths exceed $t - G_t$, while $N_{D_t} - 1$ is the smaller number of such excursions whose lengths exceed the length $D_t - G_t$ of the excursion straddling time t, for G_t and D_t defined in (69). Formula (119) is a consequence of the following analog of (117), established in [60]

$$\begin{aligned} &E\left[1(N_{D_t} = n) f\left(\frac{V_1(D_t)}{D_t}, \frac{V_2(D_t)}{D_t} \ldots\right)\right] \\ &\qquad = E_{\alpha,0}[-\alpha \log(1 - V_n) f(V_1, V_2, \ldots)], \end{aligned} \tag{120}$$

which for $f = 1$ gives

$$P(N_{D_t} = n) = E_{\alpha,0}[-\alpha \log(1 - V_n)] = \alpha \sum_{p=1}^\infty \frac{1}{p} E_{\alpha,0}[V_n^p]. \tag{121}$$

Evaluating $E_{\alpha,0}[V_n^p]$ using (49) now yields (119). Using (121) and (50) we obtain the following asymptotic formulae as $n \to \infty$:

$$P(N_{D_t} = n) \sim \alpha P(N_t = n) \sim \frac{\alpha\Gamma(1/\alpha + 1)}{\Gamma(1-\alpha)^{1/\alpha}} \frac{1}{n^{1/\alpha}}, \tag{122}$$

where $a(n) \sim b(n)$ means $a(n)/b(n) \to 1$ as $n \to \infty$. To illustrate, in the Brownian case ($\alpha = \frac{1}{2}$) the numerical values in Table 1 were obtained using a four line *Mathematica* program which evaluated the integrals (118) and (119) numerically after definition of ϕ_α and ψ_α in terms of *Mathematica*'s incomplete gamma function. The numerical values for N_{D_t} agree with those of Scheffer [62]. The asymptotic formulae as $n \to \infty$ are read from (122). For $n = 4$ the asymptotic formula gives the approximations 0.0398 and 0.0199, which are already very close to the values of $P(N_t = 4)$ and $P(N_{D_t} = 4)$ shown in Table 1.

A simplified approach to (118) and (119), which gives a probabilistic interpretation of the integrals in these formulae, can be made as follows. Let T be an exponential variable with rate 1 independent of the subordinator (τ_s). It is clear by scaling that N_t for each t has the same distribution as N_T, so it is enough to establish the formulae with T instead of t. By consideration of a Poisson process of marked excursions as in Section 3 of [59], it is found that $T - G_T$ has gamma$(1-\alpha)$ distribution, and given $T - G_T = z$ that N_T has geometric distribution with parameter $p_\alpha(z)$ as in (115). That is to say,

$$P(T - G_T \in dz, N_T = n) = \frac{1}{\Gamma(1-\alpha)} z^{-\alpha} e^{-z}\, dz\, (1 - p_\alpha(z))^{n-1} p_\alpha(z), \tag{123}$$

which gives a natural disintegration of (118) with t replaced by T. A similar argument with $D_T - G_T$ instead of $T - G_T$ yields

$$\begin{aligned} &P(D_T - G_T \in dz, N_{D_T} = n) \\ &\quad = \frac{\alpha}{\Gamma(1-\alpha)} z^{-\alpha}(1 - e^{-z})\, dz (1 - p_\alpha(z))^{n-1} p_\alpha(z), \end{aligned} \tag{124}$$

which is the corresponding disintegration of (119). To summarize, the distributions of N_t and N_{D_t} are two different integral mixtures of geometric(p) distributions on $\{1, 2, \ldots\}$; the mixing distribution is that of $p_\alpha(T - G_T)$ in the case of N_t, and that of $p_\alpha(D_T - G_T)$ in the case of N_{D_t}.

TABLE 1
Distribution of N_t and N_{D_t} for Brownian motion

n	1	2	3	4	...	$\to\infty$
$P(N_t = n)$	0.6265	0.1430	0.0630	0.0356	$\cdots$	$\sim 2/(\pi n^2)$
$P(N_{D_t} = n)$	0.8003	0.0812	0.0334	0.0185	$\cdots$	$\sim 1/(\pi n^2)$

6.4. *Interpretation in the bridge case $\theta = \alpha$.* In the case $\theta = \alpha$, corresponding to a Brownian or Bessel bridge, the distribution of N described in (114) can be understood as follows. Starting with a $(2-2\alpha)$-dimensional Bessel bridge of length 1, whose ranked excursion lengths are $V_1 > V_2 > \cdots$, let U be uniform on $[0,1]$ independent of the bridge and let $V_N = D_U - G_U$ be the length of the excursion interval (G_U, D_U) that contains time U. So V_N is a length-biased pick from the sequence of lengths (V_n). Then, as shown in Aldous and Pitman [3] for $\alpha = 1/2$, and in [52] for $0 < \alpha < 1$, the joint distribution of $(G_U, D_U - G_U, 1 - D_U)$ is Dirichlet with parameters $(\alpha, 1-\alpha, \alpha)$, and conditionally given $(G_U, D_U - G_U, 1 - D_U)$ the process B decomposes into three independent components: two bridges of lengths G_U and $1 - D_U$ and an excursion of length $D_U - G_U$. Let $V_1' > V_2' > \cdots$ denote the ranked excursion lengths up to time G_U, and let $V_1'' > V_2'' > \cdots$ denote the ranked excursion lengths derived from the interval $(D_U, 1)$. Note that the sequence $V_1 > V_2 > \cdots$ is obtained by ranking the set of lengths $V_1', V_2', \ldots, V_N, V_1'', V_2'', \ldots$ and that

$$N - 1 = N' + N'',$$

where N' is the number of i such that $V_i' > V_N$ and N'' is the number of i such that $V_i'' > V_N$. Now, if we introduce a gamma$(1+\alpha)$ random variable $Z_{1+\alpha}$ independent of the bridge, then $Z_{1+\alpha}G_U$, $Z_{1+\alpha}V_N$ and $Z_{1+\alpha}(1-D_U)$ are three independent gamma variables with parameters α, $1-\alpha$ and α, respectively, and the three random components $Z_{1+\alpha}(V_1', V_2', \ldots)$, $Z_{1+\alpha}(V_1'', V_2'', \ldots)$ and $Z_{1+\alpha}V_N$ are mutually independent. Moreover, the two infinite sequences are identically distributed and the joint law of either of these sequences with $Z_{1+\alpha}V_N$ is identical to the joint law of $(V_1(G_T), V_2(G_T), \ldots)$ with $T - G_T$ as considered in the previous section for an unconditioned Bessel process and an independent standard exponential variable T. It now follows from the previous discussion that the formula $N - 1 = N' + N''$ presents $N - 1$ as the sum of two random variables which given $Z_{1+\alpha}V_N = z$ are i.i.d. geometric with parameter $p_\alpha(z)$. Thus we recover the result (114) in the bridge case $\theta = \alpha$.

7. The Markov chain derived from PD$(\boldsymbol{\alpha}, \boldsymbol{\theta})$.

Starting from any ranked sequence of random variables $V_1 > V_2 > \cdots > 0$ with $\sum_n V_n = 1$, define new variables R_n and Y_n as in (21) and (45). Note the relations (23) and (46), which allow any one of the sequences (V_n), (Y_n) and (R_n) to be recovered from any of the others. Note also the relations

$$Y_n = (1 + R_n + R_n R_{n+1} + \cdots)^{-1} = \frac{Y_{n+1}}{Y_{n+1} + R_n}; \qquad R_n = \frac{Y_{n+1}(1 - Y_n)}{Y_n} \tag{125}$$

and the a priori constraints

$$0 < R_n < 1, \qquad 1 + R_1 + R_1 R_2 + \cdots < \infty, \qquad 0 < Y_{n+1} < Y_n/(1 - Y_n). \tag{126}$$

7.1. *The cases of* PD$(\alpha, 0)$ *and* PD$(0, \theta)$. The following elementary proposition was suggested by results of Vervaat [69, 70] and Vershik [65].

PROPOSITION 37. *Suppose that $R_1, R_2, \ldots$ are independent, and satisfy* (126) *a.s. Then (Y_n) is a Markov chain, typically with inhomogeneous transition probabilities. If the R_n are identically distributed, then (Y_n) is stationary, with homogeneous transition probabilities. If R_n has density*

$$P(R_n \in dr) = f_n(r)\,dr, \tag{127}$$

then (Y_n) has co-transition probabilities

$$\frac{P(Y_n \in dy_n | Y_{n+1} = y_{n+1})}{dy_n} = 1\left(0 < y_{n+1} < \frac{y_n}{\bar{y}_n}\right) f_n\left(\frac{y_{n+1}\bar{y}_n}{y_n}\right)\frac{y_{n+1}}{y_n^2}, \tag{128}$$

where $\bar{y}_n = 1 - y_n$.

PROOF. Since from (125), Y_{n+k} is a function of $R_{n+1}, R_{n+2}, \ldots,$ it is immediate that $Y_n = Y_{n+1}/(Y_{n+1} + R_n)$ is conditionally independent of $Y_{n+1}, Y_{n+2}, \ldots$ given Y_{n+1}. This yields the Markov property in reverse time. The formula for the co-transition probabilities is immediate by change of variable. Clearly, (Y_n) is stationary if (R_n) is i.i.d. □

Recall that $P_{\alpha,\theta}$ governs (V_n) according to the PD(α, θ) distribution. According to Theorem 8, under $P_{\alpha,0}$ for $0 < \alpha < 1$, the R_n are independent with beta$(n\alpha, 1)$ distributions. Thus Proposition 37 implies that under $P_{\alpha,0}$ the sequence (Y_n) is Markov with inhomogeneous co-transition probabilities which can be read from the proposition. The transition probabilities in the forward direction can then be written down using Bayes' rule, in terms of the density functions $p_{\alpha,0,n}(u)$, where for general (α, θ) we define

$$p_{\alpha,\theta,n}(u) = P_{\alpha,\theta}(V_n \in du)/du. \tag{129}$$

These densities are fairly complicated however. See Section 8.1.

This result under PD$(\alpha, 0)$ for $0 < \alpha < 1$ is analogous to the following result of Vershik and Shmidt [67] and Ignatov [34]: under PD$(0, \theta)$ for $\theta > 0$, the sequence (Y_n) is Markov with homogeneous transition probabilities

$$\frac{P_{0,\theta}(Y_{n+1} \in dy | Y_n = x)}{dy} = 1\left(0 < y < \frac{x}{\bar{x}} \wedge 1\right)\theta x^{-1}\bar{x}^{\theta-1}\,\frac{p_{0,\theta,1}(y)}{p_{0,\theta,1}(x)}. \tag{130}$$

While in the PD$(0, \theta)$ case the transition probabilities of the chain (Y_n) are homogeneous, the chain is not stationary. According to [67, 34], the stationary probability density for this chain is given by

$$p^*_{0,\theta}(x) = K^{-1}_{0,\theta}x^{-\theta}p_{0,\theta,1}(x), \tag{131}$$

where $K_{0,\theta}$ is a normalization constant. As shown by Ignatov [34], results of Vervaat [69] and Watterson [72] imply that

$$K_{0,\theta} = \Gamma(\theta+1)e^{\theta\gamma}, \tag{132}$$

where $\gamma = -\Gamma'(1) = 0.5771\cdots$ is Euler's constant, and that

$$p^*_{0,\theta}(x) = P^*_{0,\theta}(V_1 \in dx)/dx, \tag{133}$$

where $P^*_{0,\theta}$ makes (R_n) a sequence of i.i.d. beta$(\theta, 1)$ random variables and

$$V_1 = (1 + R_1 + R_1R_2 + \cdots + R_1R_2R_3 + \cdots)^{-1}. \tag{134}$$

The densities $p_{0,\theta,1}(x)$ and $p^*_{0,\theta}(x)$ are then determined by the $P^*_{0,\theta}$ distribution of $\Sigma_1 := (1 - V_1)/V_1$, which is the infinitely divisible law with Laplace transform

$$E^*_{0,\theta}[\exp(-\lambda\Sigma_1)] = \exp\left(-\theta\int_0^1 dx\frac{(1-e^{-\lambda x})}{x}\right). \tag{135}$$

Most of these results were obtained earlier in the special case $\theta = 1$, which arises in applications to combinatorics and number theory (see [13], [63], [24], [6], [65–67] and [15]).

It is easily verified using Proposition 37 that $P^*_{0,\theta}$ makes (Y_n) a stationary Markov chain with the same homogeneous transition probabilities as those displayed in (130) under $P_{0,\theta}$. Consequently, the above results are largely summarized by the following identity: for all positive product measurable functions f,

$$E_{0,\theta}[f(Y_1, Y_2, \ldots)] = K_{0,\theta}E^*_{0,\theta}[Y_1^\theta f(Y_1, Y_2, \ldots)]. \tag{136}$$

Note that since $V_1 = Y_1$ and the (V_n) sequence can be recovered from the (Y_n) sequence and vice versa, formula (136) holds just as well with Y_n replaced everywhere by V_n. The same is true of formula (137) below.

7.2. *Extension to* PD(α, θ). The following theorem, which is an amplification of Theorem 15, generalizes the entire collection of results described in the previous section to the full two-parameter family PD(α, θ).

THEOREM 38. *Let sequences (V_n), (R_n) and (Y_n) be related by* (21), (45) and (125). *For $0 \le \alpha < 1$, $\theta > -\alpha$, let $P_{\alpha,\theta}$ govern (V_n) with* PD(α, θ) *distribution and let $P^*_{\alpha,\theta}$ govern $(R_1, R_2, \ldots)$ as a sequence of independent random variables, such that R_n has* beta$(\theta + n\alpha, 1)$ *distribution. Then:*

(i) *for every product measurable function f,*

$$E_{\alpha,\theta}[f(Y_1, Y_2, \ldots)] = K_{\alpha,\theta}E^*_{\alpha,\theta}[Y_1^\theta f(Y_1, Y_2, \ldots)], \tag{137}$$

where $K_{0,\theta}$ is given in (132) *and*

$$K_{\alpha,\theta} = \Gamma(\theta+1)\Gamma(1-\alpha)^{\theta/\alpha} \qquad (0 < \alpha < 1, \theta > -\alpha). \tag{138}$$

(ii) *Both $P = P_{\alpha,\theta}$ and $P = P^*_{\alpha,\theta}$ govern (Y_n) as a Markov chain with the same forward transition probabilities, given by* (130) *for $\alpha = 0$ and as follows for $0 < \alpha < 1$:*

$$\frac{P(Y_{n+1} \in dy_{n+1} | Y_n = y_n)}{dy_{n+1}} = y_n^{-\alpha-1}(1-y_n)^{n\alpha+\theta-1}\frac{r(\alpha, \theta+n\alpha, y_{n+1})}{r(\alpha, \theta+n\alpha-\alpha, y_n)} \tag{139}$$

for $0 < y_n < 1$, $0 < y_{n+1} < y_n/(1-y_n)$ *and* 0 *otherwise, where*

$$r(\alpha, \theta, y)dy = \Gamma(\theta/\alpha + 1) y^\theta P^*_{\alpha,\theta}(V_1 \in dy) = C^{-1}_{\alpha,\theta} P_{\alpha,\theta}(V_1 \in dy) \tag{140}$$

for $C_{\alpha,\theta}$ *as in* (43) *and* $V_1 = Y_1$.

(iii) *The* $P^*_{\alpha,\theta}$ *distribution of* $\Sigma_1 := (1 - V_1)/V_1$ *is infinitely divisible, with Laplace transform given for* $\alpha = 0$, $\theta > 0$ *by* (135), *and for* $0 < \alpha < 1$, $\theta > -\alpha$ *by*

$$E^*_{\alpha,\theta}[\exp(-\lambda\Sigma_1)] = \left(\frac{1}{\psi_\alpha(\lambda)}\right)^{\theta/\alpha+1} \tag{141}$$

for ψ_α *as in* (34).

REMARK 39. For $0 < \alpha < 1$, the function $r(\alpha, \theta, y)$ is determined by the first equality in (140) and the Laplace transform (141). The last expression in (140) and Proposition 47 in the next section yield alternative formulae for $r(\alpha, \theta, y)$. For $\alpha = 0$, the chain (Y_n) is stationary and homogeneous under $P^*_{0,\theta}$, whereas in the case $0 < \alpha < 1$ the chain is nonhomogeneous and the distribution of Y_n depends on n. See Section 7.3 below regarding the asymptotic distribution of Y_n as $n \to \infty$.

REMARK 40. Since the results for $\alpha = 0$ are known, we shall assume for the proof that $0 < \alpha < 1$. We note however that the results for $\alpha = 0$ can be recovered by passage to the limit as $\alpha \downarrow 0$ for fixed θ, using (106).

PROOF OF THEOREM 38. Let $0 < \alpha < 1$.

(i) From the absolute continuity relation (42), for all measurable $f \geq 0$,

$$E_{\alpha,\theta}[f(Y_1, Y_2, \ldots)] = C_{\alpha,\theta} E_{\alpha,0}[L^{\theta/\alpha} f(Y_1, Y_2, \ldots)], \tag{142}$$

where L is the local time variable, which can be expressed from (24) as

$$L = Y_1^\alpha \lim_{n\to\infty} n(R_1 \cdots R_n)^\alpha \qquad (P_{\alpha,\theta} \text{ a.s., for all } \theta > -\alpha). \tag{143}$$

On the other hand, since both $P^*_{\alpha,\theta}$ and $P_{\alpha,0}$ make $R_1, \ldots, R_n$ a sequence of independent beta variables, calculating the ratio of the two product densities gives

$$\begin{aligned} &E^*_{\alpha,\theta}[f(R_1, \ldots, R_n)] \\ &\quad = \frac{\Gamma(\theta/\alpha + n + 1)}{\Gamma(\theta/\alpha + 1)\Gamma(n+1)} E_{\alpha,0}[(R_1 \cdots R_n)^\theta f(R_1, \ldots, R_n)]. \end{aligned} \tag{144}$$

Passage to the limit as $n \to \infty$, using $\Gamma(\theta/\alpha+n+1)/\Gamma(n+1) \sim n^{\theta/\alpha}$, martingale convergence and (143) yields

$$E^*_{\alpha,\theta}[f(R_1, R_2, \ldots)] = \Gamma(\theta/\alpha + 1)^{-1} E_{\alpha,0}[L^{\theta/\alpha} Y_1^{-\theta} f(R_1, R_2, \ldots)], \tag{145}$$

a formula which holds just as well with $f(Y_1, Y_2, \ldots)$ instead of $f(R_1, R_2, \ldots)$, due to (125). Comparison of (142) and (145) yields (137).

(ii) According to Proposition 37, (Y_n) is a Markov chain under $P^*_{\alpha,\theta}$ with transition probabilities which can be read from (128) and the prescribed beta density of R_n, which is $f_n(x) = (\theta + n\alpha)x^{\theta+n\alpha-1}$ for $0 < x < 1$. Bayes' rule then yields the forward transition probabilities of the form (139), for $r(\alpha, \theta, y)$ defined by the first equality in (140), after using the formula

$$P^*_{\alpha,\theta}(Y_n \in dy) = P^*_{\alpha,\theta+(n-1)\alpha}(Y_1 \in dy). \tag{146}$$

This follows from (125), since by definition the $P^*_{\alpha,\theta}$ distribution of R_n, $R_{n+1}, \dots$ is the $P^*_{\alpha,\theta+(n-1)\alpha}$ distribution of $R_1, R_2, \dots$. The second equality in (140) for $r(\alpha, \theta, y)$ is immediate from (137) and the formula (30) for $C_{\alpha,\theta}$ in (42). Since the density factor $dP^*_{\alpha,\theta}/dP_{\alpha,\theta} = K_{\alpha,\theta}Y_1^{\theta}$ is a function of Y_1, it is clear without further calculation that (Y_n) must be Markov under $P_{\alpha,\theta}$ with the same transition probabilities as under $P^*_{\alpha,\theta}$.

(iii) To obtain the formula (141) for the Laplace transform of $\Sigma_1 := (1 - V_1)/V_1$ use (145) to compute

$$E^*_{\alpha,\theta}[\exp(-\lambda\Sigma_1)] = \Gamma(\theta/\alpha + 1)^{-1} E_{\alpha,0}[X_1^{\theta/\alpha}\exp(-\lambda\Sigma_1)], \tag{147}$$

where $X_1 = LV_1^{-\alpha}$ has exponential distribution with rate 1. However, from (68),

$$E_{\alpha,0}[\exp(-\lambda\Sigma_1)|X_1] = \exp[-X_1(\psi_\alpha(\lambda) - 1)]$$

and using this expression in (147) yields (141). □

Immediately from the above theorem, we derive the formula of the following corollary, which extends formulae of Vershik and Shmidt [67], and Ignatov [34] in the case $\alpha = 0$. The Markov property of (Y_n) under $P_{\alpha,\theta}$ is evident by inspection of this formula. This formula can also be derived by suitable changes of variables and integration from Proposition 47, after changing variables and integrating out t. Combined with Proposition 37, this gives an alternative approach to the previous theorem.

COROLLARY 41. *The $P_{\alpha,\theta}$ joint density of $Y_1, \dots, Y_n$ is given by the formula*

$$\begin{aligned} &P_{\alpha,\theta}(Y_1 \in dy_1, \dots, Y_n \in dy_n)/\Pi_{i=1}^n dy_i \\ &\quad = C_{\alpha,\theta}\alpha^{n-1}\Pi_{i=1}^{n-1}[y_i^{-\alpha-1}(1-y_i)^{i\alpha+\theta-1}1(y_{i+1} < y_i/(1-y_i))] \\ &\qquad \times r(\alpha, n\alpha - \alpha + \theta, y_n) \end{aligned}$$

for $r(\alpha, \theta, y)$ defined by (140).

REMARK 42. Since $P_{\alpha,0} = P^*_{\alpha,0}$ for all $0 < \alpha < 1$, the special case $0 < \alpha < 1$, $\theta = 0$ of formula (146) allows computation of the $P_{\alpha,0}$ distribution of Y_n:

$$P_{\alpha,0}(Y_n \in dy) = \frac{1}{(n-1)!} y^{-(n-1)\alpha} r(\alpha, n\alpha - \alpha, y)\, dy. \tag{148}$$

This result can also be read from formula (81). In particular, the moments of Y_n derived from PD$(\alpha, 0)$ are given by the expression

$$E_{\alpha,0} Y_n^p = \frac{1}{(n-1)!} E_{\alpha,0}[V_1^{p-(n-1)\alpha} L^{n-1}], \tag{149}$$

which can be evaluated using (85).

REMARK 43. Note that if $(\tilde{Y}_n)$ are the independent factors as in (4) derived from the size-biased permutation $(\tilde{V}_n)$ of a PD(α, θ) sequence (V_n), then for each $k = 1, 2 \ldots$ the sequence $(\tilde{Y}_{n+k}, n = 1, 2, \ldots)$ has the same distribution as the independent factors derived similarly from the size-biased presentation of PD$(\alpha, \theta + k\alpha)$. On the other hand, the sequence $(Y_{n+k}, n = 1, 2, \ldots)$ is Markovian with the same sequence of inhomogeneous transition probabilities as (Y_n) derived from PD$(\alpha, \theta + k\alpha)$, but the initial distribution is different. This distinction appears already for $\alpha = 0$: then (Y_n) has stationary transition probabilities, but the distribution of Y_n varies with n, only approaching the stationary distribution in the limit as $n \to \infty$.

To illustrate by a concrete example, $(Y_2, Y_3, \ldots)$ derived from excursions of an unconditioned Bessel process is a Markov chain with exactly the same inhomogeneous transition function as $(Y_1, Y_2, \ldots)$ derived from the corresponding bridge. However Y_2 for the unconditioned process does not have the same law as Y_1 for the bridge.

7.3. *Asymptotic behavior of the* PD(α, θ) *chain.* It was shown by Vershik and Shmidt [67] for $\theta = 1$ and Ignatov [34] for general $\theta > 0$ that the $P_{0,\theta}$ distribution of Y_n converges to the stationary distribution (131) of the Markov chain. For $0 < \alpha < 1$, $\theta > -\alpha$, the asymptotic behavior of the distribution of Y_n can be derived as follows from the relation $Y_n = 1/(1+\Sigma_n)$ and the description of the $P_{\alpha,0}$ distribution of Σ_n provided by Proposition 11(ii). According to that proposition, under $P_{\alpha,0}$ the random variable Σ_n is the sum of n independent copies of Σ_1, which has finite moments of all orders, obtained by successive differentiations of its Laplace transform (37). In particular

$$E_{\alpha,0}(\Sigma_1) = \frac{\alpha}{1-\alpha}$$

and a strong law of large numbers implies that

$$\frac{\Sigma_n}{n} \to \frac{\alpha}{1-\alpha}, \qquad P_{\alpha,0} \text{ a.s.},$$

hence also $P_{\alpha,\theta}$ a.s. for all $\theta > -\alpha$ by Proposition 14. Similarly, the central limit theorem implies that the $P_{\alpha,0}$ distribution of

$$\sqrt{n}\left(\frac{\Sigma_n}{n} - \frac{\alpha}{1-\alpha}\right)$$

converges to the normal distribution with mean 0 and variance

$$Var_{\alpha,0}(\Sigma_1) = \frac{\alpha}{(2-\alpha)(1-\alpha)^2}.$$

A standard argument shows that this limit law under $P_{\alpha,0}$ is mixing in the sense of [2]. That is to say, the same limit distribution is obtained after a change of measure to any distribution Q that is absolutely continuous with respect to $P_{\alpha,0}$, in particular, for $Q = P_{\alpha,\theta}$ for all $\theta > -\alpha$. Translating these results in terms of $Y_n = 1/(1+\Sigma_n)$ yields the following proposition.

PROPOSITION 44. *Under $P_{\alpha,\theta}$ for all $0 < \alpha < 1$ and $\theta > -\alpha$,*

$$nY_n \to \frac{1-\alpha}{\alpha} \quad \text{a.s.} \tag{150}$$

and the distribution of

$$\sqrt{n}\left(nY_n - \frac{1-\alpha}{\alpha}\right) \tag{151}$$

converges to the normal distribution with mean 0 and variance $\alpha^{-2}(2-\alpha)^{-2}$.

These asymptotics for Y_n may be compared with the corresponding behavior of the independent factors $(\tilde{Y}_n)$ as in (4). From the beta$(1-\alpha, \theta+n\alpha)$ distribution of $\tilde{Y}_n$ under $P_{\alpha,\theta}$, one gets

$$E_{\alpha,\theta}(\tilde{Y}_n) = \frac{1-\alpha}{1+\theta+(n-1)\alpha}.$$

For $0 < \alpha < 1$, $\theta > -\alpha$, this makes

$$E_{\alpha,\theta}(n\tilde{Y}_n) \to \frac{1-\alpha}{\alpha} \quad \text{as } n \to \infty.$$

More precisely, the asymptotic distribution of $\alpha n \tilde{Y}_n$ is gamma$(1-\alpha)$. So Y_n and $\tilde{Y}_n$ are both of order $1/n$ for large n, their means are asymptotically the same, but their asymptotic distributions are different.

8. Some results for a general subordinator.

We collect in this section some results regarding interval lengths $V_n(t)$ derived for a general subordinator (τ_s) as in Section 1.2, which in the stable and gamma cases are related to PD(α, θ).

8.1. *Perman's formula.*

Let $\Delta_1 \geq \Delta_2 \geq \cdots$ be the ranked jumps up to time 1 of a drift-free subordinator $(\tau_s, s \geq 0)$. Put $V_n = \Delta_n/\tau_1$. Perman [50] found a formula for the $(n+1)$-dimensional joint density

$$p_n(t, v_1, \ldots, v_n) = P(\tau_1 \in dt, V_1 \in dv_1, \ldots, V_n \in dv_n)/dt\, dv_1 \cdots dv_n \tag{152}$$

assuming the Lévy measure Λ of (τ_s) has a density h with respect to Lebesgue measure on $(0, \infty)$. Perman's formula is as follows. For $n \geq 2$,

$$p_n(t, v_1, v_2, \ldots, v_n) = \frac{t^{n-1}h(tv_1)h(tv_2)\cdots h(tv_{n-1})}{\tilde{v}_n} p_1\left(t\tilde{v}_n, \frac{v_n}{\tilde{v}_n}\right) \tag{153}$$

for $t > 0$ and $0 < v_1 < v_2 < \cdots < v_n < 1$, $\sum_i v_i < 1$, where

$$\tilde{v}_n = 1 - v_1 - v_2 - \cdots - v_{n-1}$$

and

$$p_1(t, v) = P(\tau_1 \in dt, V_1 \in dv)/dt\,dv \tag{154}$$

is the unique solution of the integral equation

$$p_1(t, v) = th(tv) \int_0^{v/(1-v)\wedge 1} p_1(t(1-v), u)\,du \tag{155}$$

for $t > 0$ and $v \in (0, 1)$.

PROPOSITION 45. *Let $f(t) := P(\tau_1 \in dt)/dt$ denote the density of τ_1, and define a sequence of nonnegative functions $f_n(t, u)$, $t > 0$, $0 < u < 1$, inductively as*

$$f_1(t, u) = th(tu) f(t\bar{u}), \tag{156}$$

where $\bar{u} = 1 - u$ and for $n = 1, 2, \ldots,$

$$f_{n+1}(t, u) = 1(u \le 1/n) th(tu) \int_{u/\bar{u}}^1 dv f_n(t\bar{u}, v). \tag{157}$$

The joint density $p_1(t, v)$ appearing in (154) *and* (153) *is given by the formula*

$$p_1(t, v) = \sum_1^\infty (-1)^{n+1} f_n(t, v), \tag{158}$$

where all but the first n terms of the sum are zero if $v > 1/(n+1)$.

PROOF. This is straightforward by induction on n, using Perman's integral equation (155).

REMARK 46. Integrating formula (158) from u to 1 gives a series expansion for $P(V_1 > u, \tau_1 \in dt)$. It can be shown by induction that this series is identical to that obtained by Perman by a different method in formula (8) of [50].

Suppose for the rest of this section that (τ_s) is a stable subordinator of index α, as in (12). Then the density $h(x)$ of the Lévy measure is

$$h(x) = \alpha C x^{-\alpha-1} \qquad (x > 0) \tag{159}$$

and from (30) the density $f_\alpha(t)$ of τ_1 is characterized by its negative moments via the following formula: for all real $\theta > -\alpha$,

$$\int_0^\infty t^{-\theta} f_\alpha(t)\,dt = E(\tau_1^{-\theta}) = \frac{1}{C^{\theta/\alpha} C_{\alpha,\theta}} = \frac{\Gamma(\theta/\alpha + 1)}{\Gamma(\theta + 1)} \frac{1}{(C\Gamma(1-\alpha))^{\theta/\alpha}}. \tag{160}$$

PROPOSITION 47. *Let* (V_n) *have* PD$(\alpha, 0)$ *distribution and let* Σ *be defined as in* (25), *so* Σ *is the sum of the points* Δ_n *of the* PRM Λ_α *derived from* (V_n). *Then the joint density of* $(\Sigma, V_1, \dots, V_n)$ *is the function* $p_n(t, v_1, v_2, \dots, v_n)$ *given by Perman's formula* (153) *with* $h(x)$ *defined by* (159) *and* $p_1(t, v)$ *derived as in Proposition* 45 *from* $f(x) = f_\alpha(x)$ *defined by* (160). *For* (V_n) *with* PD(α, θ) *distribution, for* $0 < \alpha < 1$, $\theta > -\alpha$, *the corresponding joint density is* $c_{\alpha,\theta}\, t^{-\theta} p_n(t, v_1, v_2, \dots, v_n)$, *where* $c_{\alpha,\theta} = C^{\theta/\alpha} C_{\alpha,\theta}$.

PROOF. This is an immediate consequence of Propositions 10, 45 and 14. □

Integrating out t in the above $(n+1)$-dimensional joint density gives an expression for the n-dimensional joint density of $(V_1, \dots, V_n)$ for a PD(α, θ) distributed sequence (V_n). In particular, for $n = 1$ we obtain Proposition 20 as follows:

PROOF OF PROPOSITION 20. Proposition 47 combined with Proposition 45 yields formula (53) with the nth term of the sum replaced by the expression $(-1)^{n+1} c_{\alpha,\theta} \int_0^\infty t^{-\theta} f_{n,\alpha}(t, u)\, dt$, where $f_{n,\alpha}(t, u)$ is the $f_n(t, u)$ defined inductively by Proposition 45 starting from $f(t) = f_\alpha(t)$ as in (160). Chasing these definitions yields the expression (54) by making a suitable change of variable to simplify the integral with respect to t using (160). □

8.2. *Laplace transforms for some infinite products.* Let $V_n(T)$ be derived as in (7) from the closed range Z of a subordinator (τ_s) with Lévy measure Λ as in (8). The formulae of the following proposition serve to characterize the laws of the sequences $(V_n(s))$ and $(V_n(\tau_t)/\tau_t)$ for all $s > 0$ and $t > 0$. A formula like (161) involving just $V_1(s)$ appears as Theorem 2.1 of Knight [39]. See also formula (76) of Kingman [38] for an expression similar to (163) related to $V_1(\tau_t)/\tau_t$.

PROPOSITION 48. *For each measurable function* $g\colon (0, \infty) \to [0, 1]$ *such that* $\int_0^\infty \Lambda(dv)(1 - g(v)) < \infty$ *and* $\lambda \ge 0$,

$$\int_0^\infty ds\, e^{-\lambda s} E\biggl[\prod_n g(V_n(s))\biggr] = \frac{\int_0^\infty du e^{-\lambda u} \Lambda(u, \infty) g(u)}{\int_0^\infty \Lambda(dv)(1 - e^{-\lambda v} g(v))}, \tag{161}$$

$$\int_0^\infty ds \exp(-\lambda s) E\biggl[\prod_n g\biggl(\frac{s V_n(\tau_t)}{\tau_t}\biggr)\biggr] \tag{162}$$

$$\begin{aligned} &= \int_0^\infty du \biggl(t \int_0^\infty \Lambda(dv) \exp(-\lambda u v) g(uv) v\biggr) \\ &\quad \times \exp\biggl(-t \int_0^\infty \Lambda(dw)(1 - \exp(-\lambda u w) g(uw))\biggr). \end{aligned} \tag{163}$$

PROOF. By considering these identities with $e^{-\lambda s}g(s)$ instead of $g(s)$ it is enough to prove them for $\lambda = 0$. The left-hand side of (161) then equals

$$E\left[\sum_{u>0}\int_{\tau_{u-}}^{\tau_u} ds\left(\prod_m g(V_m(\tau_{u-}))\right)g(s-\tau_{u-})\right],$$

which, using the basic compensation formula of excursion theory, equals

$$E\left[\int_0^\infty du\left(\prod_m g(V_m(\tau_{u-}))\right)\right]\left(\int_0^\infty dv\Lambda(v,\infty)g(v)\right).$$

Now (161) follows easily after evaluating the expectation above using Fubini's theorem and the formula

$$\begin{aligned} E\left[\prod_n g(V_n(\tau_{u-}))\right] &= E\left[\prod_n g(V_n(\tau_u))\right] \\ &= \exp\left(-u\int_0^\infty \Lambda(dx)(1-g(x))\right), \end{aligned} \tag{164}$$

which expresses the fact that the $V_n(\tau_u)$ are the points of a PRM $(u\Lambda)$ ([37], (3.35)). Turning to (162), the change of variables $s = u\tau_t$ allows (162) for $\lambda = 0$ to be rewritten as

$$\int_0^\infty du\, E\left[\tau_t \prod_n g(uV_n(\tau_t))\right].$$

The integrand can be evaluated using (164) with t instead of u and $g(ux)e^{-\lambda x}$ instead of $g(x)$, by differentiation with respect to λ at $\lambda = 0$. The result is (163). □

For a stable (α) subordinator with $\Lambda = \Lambda_\alpha$ as in (12), it is easily verified that the expression in (163) equals the right-hand side of the expression in (161), which proves the identity in law of the two sequences featured in Proposition 6. Note also that (164) and hence (161) can be verified also for measurable g: $(0,\infty) \to [0,\infty)$ such that $0 < \int_0^\infty \Lambda(dv)(1-g(v)) < \infty$ provided the integral is absolutely convergent. Thus we obtain the following corollary regarding the expectation of an infinite product derived from (V_n) with PD$(\alpha, 0)$ distribution.

COROLLARY 49. *For* $0 < \alpha < 1$ *and* g: $(0,\infty) \to [0,\infty)$ *such that*

$$0 < \int_0^\infty \frac{dv}{v^{\alpha+1}}(1-g(v)) < \infty \tag{165}$$

and the integral is absolutely convergent, define

$$K_g(\alpha,\lambda) := \int_0^\infty \frac{dv}{v^{\alpha+1}}(1-e^{-\lambda v}g(v)), \tag{166}$$

$$K'_g(\alpha,\lambda) := \frac{d}{d\lambda}K_g(\alpha,\lambda) = \int_0^\infty \frac{dv}{v^\alpha}e^{-\lambda v}g(v). \tag{167}$$

Then

$$\int_0^\infty ds\, e^{-\lambda s} E_{\alpha,0}\left[\prod_n g(sV_n)\right] = \frac{K'_g(\alpha,\lambda)}{\alpha K_g(\alpha,\lambda)}. \tag{168}$$

To illustrate, taking $g(x) = \exp(-\kappa x^p)$ for $\kappa > 0$ and $p > 1$ gives a double Laplace transform which determines the distribution of $\sum_n V_n^p$ for a PD$(\alpha, 0)$ distributed (V_n). Unfortunately, such transforms seem difficult to invert. For g a polynomial with nonnegative coefficients, say

$$g(x) = 1 + \sum_{j=1}^{k} a_j x^j,$$

we find that

$$K_g(\alpha,\lambda) = \frac{\Gamma(1-\alpha)}{\alpha}\lambda^\alpha - \sum_{j=1}^{k} a_j \Gamma(j-\alpha)\lambda^{\alpha-j}.$$

Hence, the Laplace transform in (168) is

$$\frac{K'_g(\alpha,\lambda)}{\alpha K_g(\alpha,\lambda)} = \frac{1}{\lambda}\left(1 + \frac{\sum_{j=1}^k j\Gamma(j-\alpha)a_j\lambda^{k-j}}{\Gamma(1-\alpha)\lambda^k - \alpha\sum_{j=1}^k \Gamma(j-\alpha)a_j\lambda^{k-j}}\right). \tag{169}$$

In particular cases, this transform can be inverted to obtain, for example,

$$E_{\alpha,0}\left[\prod_n (1 + aV_n^p)\right] = 1 + \frac{p}{\alpha}\sum_{k=1}^{\infty}\frac{1}{(pk)!}\left(\frac{\alpha\Gamma(p-\alpha)}{\Gamma(1-\alpha)}\right)^k a^k, \tag{170}$$

which for $p = 1$ and $p = 2$ becomes

$$E_{\alpha,0}\left[\prod_n \left(1 + aV_n\right)\right] = 1 + \frac{1}{\alpha}(e^{\alpha a} - 1), \tag{171}$$

$$E_{\alpha,0}\left[\prod_n \left(1 + aV_n^2\right)\right] = 1 + \frac{2}{\alpha}\left(\cosh\left(\sqrt{\alpha(1-\alpha)a} - 1\right)\right). \tag{172}$$

Examination of the coefficients of a^k on both sides of (170) shows that (170) amounts to the following identity: for all positive integers k and p,

$$E_{\alpha,0}\left[\sum_{1\le n_1<\cdots<n_k} V_{n_1}^p \cdots V_{n_k}^p\right] = \frac{p}{\alpha}\frac{1}{(pk)!}\left(\frac{\alpha\Gamma(p-\alpha)}{\Gamma(1-\alpha)}\right)^k. \tag{173}$$

This is a special case of formula (178). Taking

$$\theta = 0, \qquad n = pk, \qquad m_p = k, \qquad m_j = 0 \qquad \text{for } j \neq p,$$

in (178) and multiplying both sides by $k!$ yields (173). Also from (178) or by variations of the above argument one can read analogs of (173) and (170) for PD(α, θ) and results for other polynomials. For instance, (168) can be inverted explicitly for $g(v) = 1 + av + bv^2$.

To conclude this section, we record the following analog of Corollary 49 for PD(α, θ) instead of PD$(\alpha, 0)$.

COROLLARY 50. *For $0 < \alpha < 1$, $\theta > 0$, $\lambda > 0$ and g and $K_g(\alpha, \lambda)$ as in Corollary 49,*

$$\int_0^\infty ds\, e^{-\lambda s} \frac{s^{\theta-1}}{\Gamma(\theta)} E_{\alpha,\theta}\left[\prod_n g(sV_n)\right] = \left(\frac{\Gamma(1-\alpha)}{\alpha K_g(\alpha,\lambda)}\right)^{\theta/\alpha}. \tag{174}$$

PROOF. This can be obtained from the previous results using formula (44), but we prefer the following derivation starting from Proposition 21. Replacing $g(v)$ by $e^{v-\lambda v} g(v)$, it suffices to establish the formula for $\lambda = 1$. Let $V_n(T)$ be derived from (τ_s) and $T = \tau(S_{\alpha,\theta})$ as in Proposition 21. By application of that Proposition, $E[\prod_n g(V_n(T))]$ equals the left-hand side of (174) for $\lambda = 1$. However, evaluating this expectation by conditioning on $S_{\alpha,\theta}$ and using (164) yields the right-hand side of (174) for $\lambda = 1$. □

APPENDIX

Here, we mention some known results which provide motivation for the definition and study of PD(α, θ).

A.1. *The finite Poisson–Dirichlet distribution.* If the convention is made that the beta(a, b) distribution is a unit mass at 1 for $a > 0$, $b = 0$, then for (α, θ) in the range

$$\alpha = -\kappa \quad \text{and} \quad \theta = m\kappa \quad \text{for some } \kappa > 0 \text{ and } m \in \{2, 3, \ldots\}. \tag{175}$$

Definition 1 prescribes a joint distribution of a finite sequence $(\tilde{V}_1, \ldots, \tilde{V}_m)$ with $\tilde{V}_i \geq 0$ and $\sum_{i=1}^m \tilde{V}_i = 1$. The distribution of the corresponding ranked sequence $(V_1, \ldots, V_m, 0, 0, \ldots)$ with $V_1 \geq \cdots V_m \geq 0$ and $\sum_{i=1}^m V_i = 1$ may still be called PD(α, θ). It is known that for $(\alpha, \theta) = (-\kappa, m\kappa)$ in this range, $(\tilde{V}_1, \ldots, \tilde{V}_m)$ may be constructed as the size-biased permutation of $(W_1, \ldots, W_m)$, where $(W_1, \ldots, W_m)$ has symmetric Dirichlet distribution obtained by setting $W_i = X_i/(X_1 + \cdots + X_m)$ for i.i.d. X_i with gamma(κ) distribution, so $(V_1, \ldots, V_m)$ can be obtained by ranking $(W_1, \ldots, W_m)$. See [37], Section A.6, for a proof and references. As shown by Kingman [38], as $\kappa = -\alpha \downarrow 0$ and $m \uparrow \infty$ for fixed $\theta = m\kappa$, PD(α, θ) converges weakly to PD$(0, \theta)$. It is easily verified that the formulae in this paper which follow directly from Proposition 2, in particular, (6) (52) and (178), hold also for (α, θ) in the range (175). See also [25] for some moment formulae for the finite Poisson–Dirichlet distribution in the vein of (51).

A.2. *The partition structure derived from* PD(α, θ). In a random sample of size n from a population with random frequencies $(V_1, V_2, \ldots)$ and a vector of nonnegative integers $(m_1, \ldots, m_n)$ with $\Sigma i m_i = n$, the probability that there are m_1 species with a single representative in the sample and m_2 species with two representatives in the sample and so on, is given by the formula

$$p(m_1, \ldots, m_n) = \frac{n!}{\prod_{i=1}^n (i!)^{m_i} m_i!} \mu(m_1, \ldots, m_n) \tag{176}$$

with

$$\mu(m_1, \dots, m_n) = E\left[\sum \prod_{i=1}^{n} \prod_{j=1}^{m_i} V^i_{n(i,\,j)}\right], \tag{177}$$

where the summation ranges over all choices of distinct $n(i, j)$ with

$$i = 1, \dots, n; \qquad j = 1, \dots, m_i.$$

See Kingman [37], where the expectation (177) is evaluated for (V_n) with PD(0, θ) distribution to obtain the formula for $p(m_1, \dots m_n)$ in this case, which is the *Ewens sampling formula* [19–21]. Proposition 9 of Pitman [54] gives the generalization of the Ewens formula for PD(α, θ), which can be stated as follows. For real numbers x and a and nonnegative integer m, let

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

and let $[x]_m = [x]_{m,1}$. Note that $[1]_m = m!$.

PROPOSITION 51 [54]. *For* (V_n) *with* PD(α, θ) *distribution,* (176) *and* (177) *hold with* $\mu(m_1, \dots, m_n) = \mu_{\alpha,\theta}(m_1, \dots, m_n)$ *given by the formula*

$$\mu_{\alpha,\theta}(m_1, \dots, m_n) = \frac{[\theta+\alpha]_{k-1,\,\alpha}}{[\theta+1]_{n-1}} \prod_{j=1}^{n} \Big([1-\alpha]_{j-1}\Big)^{m_j}. \tag{178}$$

See [52–55, 36] for various developments and applications of this formula. As a consequence of Proposition 51, the urn scheme for generating PD(0, θ) studied by various authors [8, 28, 30, 14] also admits a two-parameter generalization [54, 57], whose simple form provides another characterization of the two-parameter family [75].

REFERENCES

[1] ALDOUS, D. (1985). Exchangeability and related topics. *École d'Été de Probabilités de Saint-Flour XII. Lecture Notes in Math.* 1117. Springer, Berlin.

[2] ALDOUS, D. J. and EAGLESON, G. K. (1978). On mixing and stability of limit theorems. *Ann. Probab.* 6 325–331.

[3] ALDOUS, D. and PITMAN, J. (1994). Brownian bridge asymptotics for random mappings. *Random Structures Algorithms* 5 487–512.

[4] AROV, D. and BOBROV, A. (1960). The extreme terms of a sample and their role in the sum of independent variables. *Theory Probab. Appl.* 5 377–396.

[5] BARLOW, M., PITMAN, J. and YOR, M. (1989). Une extension multidimensionnelle de la loi de l'arc sinus. In *Séminaire de Probabilités XXIII. Lecture Notes in Math.* 1372 294–314. Springer, Berlin.

[6] BILLINGSLEY, P. (1972). On the distribution of large prime factors. *Period. Math. Hungar.* 2 283–289.

[7] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Encylopedia of Mathematics and Its Applications*: *Regular Variation*. Cambridge Univ. Press.

[8] BLACKWELL, D. and MACQUEEN, J. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* 1 353–355.

[9] BROCKWELL, P. J. and BROWN, B. M. (1978). Expansion for the positive stable laws. *Z. Wahrsch. Verw. Gebiete* 45 213–224.

[10] CHUNG, K. and ERDÖS, P. (1952). On the application of the Borel–Cantelli Lemma. *Trans. Amer. Math. Soc.* 72 179–186.

[11] CSÁKI, E., ERDÖS, P. and REVESZ, P. (1985). On the length of the longest excursion. *Z. Wahrsch. Verw. Gebiete* 68 365–382.

[12] DARLING, D. (1952). The influence of the maximum term in the addition of independent random variables. *Trans. Amer. Math. Soc.* 73 95–107.

[13] DICKMAN, K. (1930). On the frequency of numbers containing prime factors of a certain relative magnitude. *Arkiv. for Matematik Astronomi och Fysik* 22 1–14.

[14] DONNELLY, P. (1986). Partition structures, Pólya urns, the Ewens sampling formula, and the ages of alleles. *Theoret. Population Biol.* 30 271–288.

[15] DONNELLY, P. and GRIMMETT, G. (1993). On the asymptotic distribution of large prime factors. *J. London Math. Soc.* (*2*) 47 395–404.

[16] DONNELLY, P. and JOYCE, P. (1989). Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stochastic Process. Appl.* 31 89–103.

[17] ENGEN, S. (1978). *Stochastic Abundance Models with Emphasis on Biological Communities and Species Diversity*. Chapman and Hall, London.

[18] EVANS, S. and PITMAN, J. (1996). Construction of Markovian coalescents. Technical Report 465, Dept. Statistics, Univ. California, Berkeley.

[19] EWENS, W. (1972). The sampling theory of selectively neutral alleles. *Theoret. Population Biol.* 3 87–112.

[20] EWENS, W. (1988). Population genetics theory—the past and the future. In *Mathematical and Statistical Problems in Evolution* (S. Lessard, ed.). Univ. Montreal Press.

[21] EWENS, W. and TAVARÉ, S. (1995). The Ewens sampling formula. In *Multivariate Discrete Distributions* (N. S. Johnson, S. Kotz and N. Balakrishnan, eds.). Wiley, New York. To appear.

[22] FERGUSON, T. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* 1 209–230.

[23] GETOOR, R. (1963). The asymptotic distribution of the number of zero free intervals of a stable process. *Trans. Amer. Math. Soc.* 106 127–138.

[24] GONČHAROV, V. (1962). On the field of combinatory analysis. *Amer. Math. Soc. Transl.* 19 1–46.

[25] GRIFFITHS, R. C. (1979). Exact sampling distributions from the infinite neutral alleles models. *Adv. in Appl. Probab.* 11 326–354.

[26] GRIFFITHS, R. C. (1988). On the distribution of points in a Poisson Dirichlet process. *J. Appl. Probab.* 25 336–345.

[27] HANSEN, J. (1994). Order statistics for decomposable combinatorial structures. *Random Structures Algorithms* 5 517–533.

[28] HOPPE, F. M. (1984). Pólya-like urns and the Ewens sampling formula. *J. Math. Biol.* 20 91–94.

[29] HOPPE, F. M. (1986). Size-biased filtering of Poisson–Dirichlet samples with an application to partition structures in genetics. *J. Appl. Probab.* 23 1008–1012.

[30] HOPPE, F. M. (1987). The sampling theory of neutral alleles and an urn model in population genetics. *J. Math. Biol.* 25 123–159.

[31] HOROWITZ, J. (1971). A note on the arc-sine law and Markov random sets. *Ann. Math. Statist.* 42 1068–1074.

[32] HOROWITZ, J. (1972). Semilinear Markov processes, subordinators and renewal theory. *Z. Wahrsch. Verw. Gebiete* 24 167–193.

[33] HU, Y. and SHI, Z. (1995). Extreme lengths in Brownian and Bessel excursions. Preprint, Laboratoire de Probabilités, Univ. Paris VI.

[34] IGNATOV, T. (1982). On a constant arising in the theory of symmetric groups and on Poisson–Dirichlet measures. *Theory Probab. Appl.* 27 136–147.

[35] KALLENBERG, O. (1973). Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrsch. Verw. Gebiete* 27 23–36.

[36] KEROV, S. (1995). Coherent random allocations and the Ewens–Pitman formula. PDMI preprint, Steklov Math. Institute, St. Petersburg.

[37] KINGMAN, J. (1993). *Poisson Processes*. Clarendon, Oxford.

[38] KINGMAN, J. F. C. (1975). Random discrete distributions. *J. Roy. Statist. Soc. Ser. B* 37 1–22.

[39] KNIGHT, F. (1985). On the duration of the longest excursion. In *Seminar on Stochastic Processes* (E. Çinlar, K. Chung and R. Getoor, eds.) 117–148. Birkhäuser, Basel.

[40] KÜCHLER, U. and LAURITZEN, S. L. (1989). Exponential families, extreme point models and minimal space-time invariant functions for stochastic processes with stationary and independent increments. *Scand. J. Statist.* 16 237–261.

[41] LAMPERTI, J. (1958). An occupation time theorem for a class of stochastic processes. *Trans. Amer. Math. Soc.* 88 380–387.

[42] LAMPERTI, J. (1961). A contribution to renewal theory. *Proc. Amer. Math. Soc.* 12 724–731.

[43] LAMPERTI, J. (1962). An invariance principle in renewal theory. *Ann. Math. Statist.* 33 685–696.

[44] LAMPERTI, J. (1972). Semi-stable Markov processes. I. *Z. Wahrsch. Verw. Gebiete* 22 205–225.

[45] LEPAGE, R., WOODROOFE, M. and ZINN, J. (1981). Convergence to a stable distribution via order statistics. *Ann. Probab.* 9 624–632.

[46] LÉVY, P. (1939). Sur certains processus stochastiques homogènes. *Compositio Math.* 7 283–339.

[47] LUKACS, E. (1955). A characterization of the gamma distribution. *Ann. Math. Statist.* 26 319–324.

[48] MCCLOSKEY, J. W. (1965). A model for the distribution of individuals by species in an environment. Ph.D. thesis, Michigan State Univ.

[49] MOLCHANOV, S. A. and OSTROVSKI, E. (1969). Symmetric stable processes as traces of degenerate diffusion processes. *Theory Probab. Appl.* 14 128–131.

[50] PERMAN, M. (1993). Order statistics for jumps of normalized subordinators. *Stochastic Process Appl.* 46 267–281.

[51] PERMAN, M., PITMAN, J. and YOR, M. (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields* 92 21–39.

[52] PITMAN, J. (1992). Partition structures derived from Brownian motion and stable subordinators. *Bernoulli*. To appear.

[53] PITMAN, J. (1992). The two-parameter generalization of Ewens' random partition structure. Technical Report 345, Dept. Statistics, Univ. California, Berkeley.

[54] PITMAN, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* 102 145–158.

[55] PITMAN, J. (1995). Species sampling models. Unpublished manuscript.

[56] PITMAN, J. (1996). Random discrete distributions invariant under size-biased permutation. *Adv. in Appl. Probab.* 28 525–539.

[57] PITMAN, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics, Probability and Game Theory. Papers in Honor of David Blackwell*. IMS, Hayward, CA. To appear.

[58] PITMAN, J. (1996). The additive coalescent, random trees, and Brownian excursions. Unpublished manuscript.

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

[60] PITMAN, J. and YOR, M. (1996). On the relative lengths of excursions derived from a stable subordinator. *Séminaire de Probabilités XXXI*. To appear.

[61] RESNICK, S. I. (1986). Point processes, regular variation and weak convergence. *Adv. in Appl. Probab.* 18 66–138.

[62] SCHEFFER, C. (1995). The rank of the present excursion. *Stochastic Process Appl.* 55 101–118.

[63] SHEPP, L. and LLOYD, S. (1966). Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.* 121 340–357.

[64] STEPANOV, V. (1969). Limit distributions of certain characteristics of random mappings. *Theory Probab. Appl.* 14 612–626.

[65] VERSHIK, A. M. (1986). The asymptotic distribution of factorizations of natural numbers into prime divisors. *Soviet Math. Dokl.* 34 57–61.

[66] VERSHIK, A. and SHMIDT, A. (1977). Limit measures arising in the theory of groups, I. *Theory Probab. Appl.* 22 79–85.

[67] VERSHIK, A. and SHMIDT, A. (1978). Limit measures arising in the theory of symmetric groups, II. *Theory Probab. Appl.* 23 36–49.

[68] VERSHIK, A. and YOR, M. (1995). Multiplicativité du processus gamma et étude asymptotique des lois stables d'indice α, lorsque α tend vers 0. Technical Report 289, Laboratoire de Probabilités, Univ. Paris VI.

[69] VERVAAT, W. (1972). *Success Epochs in Bernoulli Trials. Math. Centre Tracts* 42. Math. Centrum, Amsterdam.

[70] VERVAAT, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. in Appl. Probab.* 11 750–783.

[71] WATANABE, S. (1995). Generalized arc-sine laws for one-dimensional diffusion processes and random walks. *Proc. Sympos. Pure Math.* 57 157–172.

[72] WATTERSON, G. A. (1976). The stationary distribution of the infinitely-many neutral alleles diffusion model. *J. Appl. Probab.* 13 639–651.

[73] WATTERSON, G. A. and GUESS, H. (1977). Is the most frequent allele the oldest? *Theoret. Population Biol.* 11 141–160.

[74] WENDEL, J. (1964). Zero-free intervals of semi-stable Markov processes. *Math. Scand.* 14 21–34.

[75] ZABELL, S. (1996). The continuum of inductive methods revisited. In *The Cosmos of Science* (J. Earman and J. Norton, eds.). Univ. Pittsburgh Press/Univ. Konstanz. To appear.

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

DEPARTMENT OF STATISTICS
367 EVANS HALL #3860
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720-3860
E-MAIL: pitman@stat.berkeley.edu