J. Appl. Probab. 1–18 (2025)
*doi:*10.1017/jpr.2025.28

INVERSE CLUSTERING OF GIBBS PARTITIONS VIA INDEPENDENT FRAGMENTATION AND DUAL DEPENDENT COAGULATION OPERATORS

MAN-WAI HO,* AND
LANCELOT F. JAMES,* ** *The Hong Kong University of Science and Technology*
JOHN W. LAU,*** *The University of Western Australia*

Abstract

Exchangeable partitions of the integers and their corresponding mass partitions on $\mathcal{P}_\infty = \{\mathbf{s} = (s_1, s_2, \ldots)\colon s_1 \geq s_2 \geq \cdots \geq 0 \text{ and } \sum_{k=1}^{\infty} s_k = 1\}$ play a vital role in combinatorial stochastic processes and their applications. In this work, we continue our focus on the class of Gibbs partitions of the integers and the corresponding stable Poisson–Kingman-distributed mass partitions generated by the normalized jumps of a stable subordinator with an index $\alpha \in (0, 1)$, subject to further mixing. This remarkable class of infinitely exchangeable random partitions is characterized by probabilities that have Gibbs (product) form. These partitions have practical applications in combinatorial stochastic processes, random tree/graph growth models, and Bayesian statistics. The most notable class consists of random partitions generated from the two-parameter Poisson–Dirichlet distribution $\mathrm{PD}(\alpha, \theta)$. While the utility of Gibbs partitions is recognized, there is limited understanding of the broader class. Here, as a continuation of our work, we address this gap by extending the dual coagulation/fragmentation results of Pitman (1999), developed for the Poisson–Dirichlet family, to all Gibbs models and their corresponding Poisson–Kingman mass partitions, creating nested families of Gibbs partitions and mass partitions. We focus primarily on fragmentation operations, identifying which classes correspond to these operations and providing significant calculations for the resulting Gibbs partitions. Furthermore, for completion, we provide definitive results for dual coagulation operations using dependent processes. We demonstrate the applicability of our results by establishing new findings for Brownian excursion partitions, Mittag-Leffler, and size-biased generalized gamma models.

Keywords: Brownian and Bessel processes; coagulation/fragmentation duality; Gibbs partitions; Poisson–Dirichlet distributions; stable Poisson–Kingman distributions

2020 Mathematics Subject Classification: Primary 62G05; 60G09
Secondary 60G57; 60E99

Received 10 April 2025; accepted 10 April 2025.

* Postal address: Department of Information Systems, Business Statistics and Operations Management, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR.
** Email address: lancelot@ust.hk
*** Postal address: UWA Centre for Applied Statistics, The University of Western Australia (M019), 35 Stirling Highway, CRAWLEY WA 6009, Australia.

1. Introduction

Infinitely exchangeable random partitions, denoted as Π_∞, of the integers $\mathbb{N} = \{1, 2, \ldots\}$, following Kingman, correspond to sampling from a discrete random distribution function representable as $F(y) = \sum_{k=1}^{\infty} P_k 1_{\{U_k \le y\}}$, where $(U_k)_{k\ge1} \overset{\text{iid}}{\sim} \text{Uniform}(0, 1)$ and are independent of the random mass partitions $(P_k)_{k\ge1} \in \mathcal{P}_\infty = \{\mathbf{s} = (s_1, s_2, \ldots) \colon s_1 \ge s_2 \ge \cdots \ge 0 \text{ and } \sum_{k=1}^{\infty} s_k = 1\}$, as described in [3, 32, 33]. These concepts play important roles in areas such as Bayesian statistics, excursion theory, population genetics, species sampling models, statistical physics, and random graph theory, where combinatorial stochastic processes frequently occur as fundamental components.

Specifically, let $\Pi_n := \{C_1, \ldots, C_{K_n}\}$ denote partitions of the integers $[n] := \{1, \ldots, n\}$, with $K_n \le n$ representing the random number of blocks and block sizes $n_j = |C_j|$ such that $\sum_{j=1}^{K_n} n_j = n$. This defines the restriction of Π_∞ to Π_n. Following Kingman's correspondence [32, 33], if $X_1, \ldots, X_n$ represents an independent and identically distributed (iid) sample from F with unique values $(\tilde{U}_1, \ldots, \tilde{U}_k) \overset{\text{iid}}{\sim} \text{Uniform}[0, 1]$ for $K_n = k$, each block of the partition can be represented as $C_j = \{i \colon X_i = \tilde{U}_j\}$ for $j = 1, \ldots, k$. The law of Π_n is expressed as $\mathbb{P}(\Pi_n := \{C_1, \ldots, C_k\}) = p(n_1, \ldots, n_k)$, where p is a symmetric function of the integers depending solely on the sizes of the blocks $(n_1, \ldots, n_k)$, and hence not on the values $\tilde{U}_j$, known as exchangeable partition probability functions (EPPFs) of the corresponding infinite partition Π_∞ of $\mathbb{N}$. Furthermore, the EPPF specifying the distribution of Π_∞ is in bijection to the distribution of a corresponding $(P_k)_{k\ge1} \in \mathcal{P}_\infty$ and the frequencies $|C_j|/n, j = 1, \ldots, K_n$, arranged in ranked order converge almost surely to $(P_k)_{k\ge1}$ as $n \to \infty$.

This work continues the research initiated in [20, 21], focusing on finer distributional details of a large class of infinitely exchangeable random partitions known as Gibbs partitions and corresponding stable Poisson–Kingman-distributed mass partitions. As established in [32, Theorem 8] and subsequently in [15, 33], this large class has the distinguishing feature that the EPPFs for Π_∞ have Gibbs or product form consistent with each restriction to Π_n. Furthermore, these partitions correspond to sampling from F where $(P_k)_{k\ge1} \sim \text{PK}_\alpha(\gamma)$ denotes α-stable Poisson–Kingman distributions with mixing (probability) measure γ on $(0, \infty)$, constructed from the normalization of the jumps of an α-stable subordinator, where $\alpha \in (0, 1)$.

Remark 1.1. As in [3], we may refer to F as a $\text{PK}_\alpha(\gamma)$-bridge, and the corresponding $\Pi_n := \{C_1, \ldots, C_{K_n}\}$ as a $\text{PK}_\alpha(\gamma)$ partition of $[n]$.

Let T_α be a positive α-stable random variable with Laplace transform $\mathbb{E}[e^{-\lambda T_\alpha}] = e^{-\lambda^\alpha}$ and density f_α. In addition, $T_\alpha^{-\alpha} \sim \text{ML}(\alpha, 0)$ denotes a variable with a Mittag-Leffler distribution with corresponding density g_α. We can set $\gamma(\mathrm{d}t)/\mathrm{d}t = h(t) f_\alpha(t)$ for a non-negative function $h(t)$, for $t \in (0, \infty)$, such that $\mathbb{E}[h(T_\alpha)] = 1$. Hence, we can write $(P_k)_{k\ge1} \sim \text{PK}_\alpha(h \cdot f_\alpha)$ in place of $\text{PK}_\alpha(\gamma)$, emphasizing the change in distribution based on the choices of $h(t)$. The case where $h(t) = 1$, for $t > 0$, yields the Poisson–Dirichlet distribution $\text{PD}(\alpha, 0) := \text{PK}_\alpha(f_\alpha)$, which is the canonical class based on a normalized stable subordinator. Sampling from the corresponding $F := F_\alpha$, a $\text{PD}(\alpha, 0)$-bridge, yields the EPPF, say p_α, of a $\text{PD}(\alpha, 0)$ (Gibbs) partition of $[n]$:

$$\mathbb{P}(\Pi_n := \{C_1, \ldots, C_k\}) = p_\alpha(n_1, \ldots, n_k) := \frac{\alpha^{k-1}\Gamma(k)}{\Gamma(n)} \prod_{j=1}^{k} (1 - \alpha)_{n_j - 1},$$

where, for any non-negative integer x, $(x)_n = x(x + 1) \cdots (x + n - 1) = \Gamma(x + n)/\Gamma(x)$ denotes the Pochhammer symbol, with more details provided in Section 2. Furthermore,

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

$\mathbb{P}_{\alpha,0}(K_n=k)=\mathbb{P}^{(n)}_{\alpha,0}(k)$ is the corresponding distribution of the number of blocks K_n, where

$$\mathbb{P}^{(n)}_{\alpha,0}(k)=\frac{\alpha^{k-1}\Gamma(k)}{\Gamma(n)}S_\alpha(n,k), \quad \text{with } S_\alpha(n,k)=\frac{1}{\alpha^k k!}\sum_{j=1}^{k}(-1)^j\binom{k}{j}(-j\alpha)_n$$

denoting the generalized Stirling number of the second kind. In general, as established in [15, 32, 33], the distribution of the $\mathrm{PK}_\alpha(h\cdot f_\alpha)$ (Gibbs) partition of $[n]$ is specified by the EPPF, say $p^{[\gamma]}_\alpha$, expressed as

$$\mathbb{P}(\Pi_n:=\{C_1,\ldots,C_k\})=p^{[\gamma]}_\alpha(n_1,\ldots,n_k)=\Psi^{[\alpha]}_{n,k}\times p_\alpha(n_1,\ldots,n_k), \tag{1.1}$$

where $\Psi^{[\alpha]}_{n,k}=\mathbb{E}[h(T_\alpha)\,|\,K_n=k]$ evaluated under the $\mathrm{PD}(\alpha,0)$ distribution of $T_\alpha\,|\,K_n=k$, as described in [20, 21]. The most distinguished and utilized class of these partition models is generated by $(P_k)_{k\geq1}\sim\mathrm{PD}(\alpha,\theta)$ for $0\leq\alpha<1$ and $\theta>-\alpha$, denoting the two-parameter Poisson–Dirichlet distribution arising for $h(t)=t^{-\theta}/\mathbb{E}[T_\alpha^{-\theta}]$. $T_{\alpha,\theta}$ is the analogue of T_α with density $f_{\alpha,\theta}(t)=t^{-\theta}f_\alpha(t)/\mathbb{E}[T_\alpha^{-\theta}]$, and $T^{-\alpha}_{\alpha,\theta}\sim\mathrm{ML}(\alpha,\theta)$ are general Mittag-Leffler variables with parameters (α,θ). The size-biased rearrangement of $(P_k)_{k\geq1}$, say $\big(\tilde{P}_k=(1-R_k)\prod_{j=1}^{k-1}R_j\big)_{k\geq1}\sim\mathrm{GEM}(\alpha,\theta)$, where $\mathrm{GEM}(\alpha,\theta)$ is the two-parameter Griffiths–Engen–McCloskey distribution, means that $1-R_k\overset{\text{ind}}{\sim}\mathrm{Beta}(1-\alpha,\theta+k\alpha)$ for $k\geq1$, and is widely used in applications. The EPPFs corresponding to $\{C_1,\ldots,C_{K_n}\}\sim\mathrm{CRP}(\alpha,\theta)-[n]$, denoting the distribution of a two-parameter Chinese restaurant process partition of $[n]$, equivalently a $\mathrm{PD}(\alpha,\theta)$ partition of $[n]$, is usually denoted as $p_{\alpha,\theta}(n_1,\ldots,n_k)$ and has the form in (1.1) with

$$\Psi^{[\alpha]}_{n,k}=\frac{\mathbb{E}\big[T_\alpha^{-\theta}\,|\,K_n=k\big]}{\mathbb{E}\big[T_\alpha^{-\theta}\big]}=\frac{\Gamma(n)\Gamma(\theta/\alpha+k)\Gamma(\theta+1)}{\Gamma(k)\Gamma(\theta/\alpha+1)\Gamma(\theta+n)}, \tag{1.2}$$

where $\mathbb{E}[T_\alpha^{-\theta}\,|\,K_n=k]=\Gamma(n)\Gamma(\theta/\alpha+k)/[\Gamma(k)\Gamma(\theta+n)]$. Furthermore, the corresponding $F:=F_{\alpha,\theta}$ are said to be Pitman–Yor processes, as coined in [22], or otherwise just $\mathrm{PD}(\alpha,\theta)$-bridges. See [3, 22, 30, 33, 37] for more details. These are extensions of results associated with the classic one-parameter Poisson–Dirichlet distribution corresponding to $\mathrm{PD}(0,\theta)$, including the Ewens sampling formula arising in population genetics; here $F:=F_{0,\theta}$ is a Dirichlet process as in [13]. See [9] for a recent comprehensive overview. Also, see [15, Theorem 12] for a precise description of the entire class of Gibbs partitions. Another popular choice is $h(t)=\mathrm{e}^{-\lambda t}\mathrm{e}^{\lambda^\alpha}$, corresponding to generalized gamma processes. However, in general, interpretation of distributions based on different choices of $h(t)$ is lacking.

1.1. Outline

In this work, in part to obtain a better understanding of various choices of $h(t)$, we tackle the formidable challenge of extending the dual coagulation/fragmentation results of [31], developed for the two-parameter Poisson–Dirichlet family, to all Gibbs models, thereby creating nested families of Gibbs partitions and mass partitions. In general, fragmentation refers to shattering/splitting of mass partitions or blocks of integer partitions, whereas coagulation refers to merging of such objects [3, 33]. Our primary focus is on fragmentation operations on $\mathcal{P}_\infty$ and their corresponding actions on Π_∞. We identify which classes correspond to these operations and provide simplified calculations for the resulting Gibbs partitions, thus providing a better indication of the distributional meaning of the resulting mixing measures. For completeness, we also describe the dual coagulation, which, unlike the Poisson–Dirichlet case in [31],

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

is based on dependent processes. See [3, 33] for more general accounts of coagulation and fragmentation operations. Readers interested in applications in statistical mechanics may refer to [3, Section 4.4.4] in relation to the Bolthausen–Sznitman coalescent [6], and the works of [31] and others.

A more detailed outline is as follows. In Section 2, we provide more background details for the $\mathrm{PK}_\alpha(h\cdot f_\alpha)$ distributions and corresponding EPPFs. In Section 3, for $0\le\beta<\alpha$ we introduce the fragmentation operator $\mathrm{FRAG}_{\alpha,-\beta}$ on $\mathcal{P}_\infty$ and establish new results for its application to $\mathrm{PK}_\beta(h\cdot f_\beta)$ mass partitions, where to be clear $h(t)$ is chosen such that $\mathbb{E}[h(T_\beta)]=1$. The $\mathrm{FRAG}_{\alpha,-\beta}$ operation then induces a new measure based on a modification of h and the density f_α. In Section 4, we establish the dual coagulation operation based on dependent processes. In Sections 5.1 and 5.2 we obtain remarkable calculations for Gibbs partitions and related identities, derived from the $\mathrm{FRAG}_{\alpha,-\beta}$ operation, which are not obvious from our results in Section 3. The remainder of our exposition focuses on applications of our results to obtain new results for non-trivial special cases of interest. In Section 5.3, we focus on the application of the fragmentation operator $\mathrm{FRAG}_{\alpha,-1/2}$ to Brownian excursion partitions [32, 33]. Section 5.4 introduces new fragmentation results in relation to the Mittag-Leffler class [21]. Section 6 demonstrates how our dual coagulation and fragmentation operations can be used to identify all the relevant laws and construct new duality results for generalized gamma models and their size-biased extensions.

For nested models primarily related to the fragmentation operator, see [2, 10]. In addition, [38] (see also [14]) applies the coagulation/fragmentation duality on the space of partitions of $[n]$ to ∞-gram natural language models. This represents an application in Bayesian statistical machine learning involving the use of inverse clustering (via $\mathrm{FRAG}_{\alpha,-\beta}$ fragmentation) and merging (via $\mathrm{PD}(\beta/\alpha,\theta/\alpha)$ coagulation) on the space of partitions of $[n]$. Related to this, [26] constructs (nested) hierarchical network/graph models using the coagulation/fragmentation operations in [31], and also [12]. This suggests that our results can potentially be applied in a similar manner to graphs constructed from generalized gamma models, which is a flexible class highlighted in [7, Section 6.3]. For some other references on Gibbs partitions and α-stable Poisson–Kingman models, see [1, 7, 11, 17, 20, 21, 28, 35]. See [18, 19] for other occurrences of the coagulation/fragmentation operators in the $\mathrm{PD}(\alpha,\theta)$ setting.

Remark 1.2. Pitman [31, Section 4] presents a proof of his duality result in Theorem 12, expressed on the space of integer partitions Π_∞, by employing calculations involving the relevant EPPFs. This establishes the corresponding results on the space of mass partitions $\mathcal{P}_\infty$. An attempt to extend his proof to more general Gibbs distributions on Π_∞ does not seem tenable. Instead, we focus on the space of mass partitions $\mathcal{P}_\infty$, facilitated by constructions in [5] that allow us to utilize the existing results of [31] more transparently and identify the resulting mixing measures.

2. α-stable Poisson–Kingman distributions and Gibbs partitions

Letting $(\Delta_k)_{k\ge1}$, with $\Delta_1>\Delta_2>\cdots$, denote the ranked jumps of an α-stable subordinator, say $\mathbf{T}_\alpha:=(T_\alpha(t)\colon t\ge0)$ of index $\alpha\in(0,1)$, with $T_\alpha(0)=0$ and $T_\alpha(1):=T_\alpha=\sum_{k=1}^{\infty}\Delta_k$ a positive stable random variable having Laplace transform $\mathbb{E}[\mathrm{e}^{-\lambda T_\alpha}]=\mathrm{e}^{-\lambda^\alpha}$ and density denoted as $f_\alpha(t)$. In addition, let $(L_t\colon t\ge0)$ denote the inverse of $\mathbf{T}_\alpha$, which corresponds to the local-time process of a strong Markov process starting at 0, for instance a Bessel process of dimension $2-2\alpha$ starting at zero. As in [36], we may use the notation $L(t)$ in place of L_t for notational convenience. The local time up to time 1 satisfies $L(1):=L_1\overset{\mathrm{d}}{=}T^{-\alpha}\sim\mathrm{ML}(\alpha,0)$, having a Mittag-Leffler distribution with density denoted as g_α. Hereafter, we will simply say $L(1) := L_1$ is the local time at 1. That is, $T_\alpha(s) = \inf\{t\colon L_t > s\}$, $s \geq 0$, is the corresponding inverse local-time process. We have the following important scaling identity (see [36]):

$$L_1 \overset{\mathrm{d}}{=} \frac{L_t}{t^\alpha} \overset{\mathrm{d}}{=} \frac{s}{[T_\alpha(s)]^\alpha} \overset{\mathrm{d}}{=} T_\alpha^{-\alpha}. \tag{2.1}$$

Then $(P_k := \Delta_k / T_\alpha)_{k\geq 1} \sim \mathrm{PD}(\alpha, 0)$ [37], a two-parameter $(\alpha, 0)$ Poisson–Dirichlet distribution, where as an example $\mathrm{PD}\big(\frac{1}{2}, 0\big)$ corresponds to the distribution of the lengths of excursions of a Brownian motion on [0,1]. Furthermore, $F(u) := F_\alpha(u) \overset{\mathrm{d}}{=} T_\alpha(u)/T_\alpha(1)$ for $u \in [0, 1]$. Another important property, relevant to species sampling or species abundance models, is that $n^{-\alpha} K_n$ converges almost surely to $T_\alpha^{-\alpha}$, where in this context $T_\alpha^{-\alpha}$ is referred to as the α-diversity of Π_∞, as in [32, 33].

To construct the larger class of distributions, let $\mathrm{PD}(\alpha \mid t)$ denote the conditional distribution of $(P_k)_{k\geq 1}$ given $T_\alpha = t$ (or $L_1 = t^{-\alpha}$) and, for a non-negative function $h(t)$ satisfying $\mathbb{E}[h(T_\alpha)] = 1$, mix $\mathrm{PD}(\alpha \mid t)$ over the probability measure $\gamma(\mathrm{d}t)/\mathrm{d}t = h(t) f_\alpha(t)$ to obtain $(P_k)_{k\geq 1} \sim \mathrm{PK}_\alpha(h \cdot f_\alpha)$, where

$$\mathrm{PK}_\alpha(h \cdot f_\alpha) = \int_0^\infty \mathrm{PD}(\alpha \mid t) h(t) f_\alpha(t)\, \mathrm{d}t = \int_0^\infty \mathrm{PD}\big(\alpha \mid s^{-1/\alpha}\big) h\big(s^{-1/\alpha}\big) g_\alpha(s)\, \mathrm{d}s.$$

It follows from [20, 21] that the conditional distribution of $T_\alpha \mid K_n = k$ equates to the distribution of the variable $Y_{\alpha,k\alpha}^{(n-k\alpha)}$, with density $f_{\alpha,k\alpha}^{(n-k\alpha)}(t)$, such that (pointwise), as in [21, (2.13), p. 323],

$$Y_{\alpha,k\alpha}^{(n-k\alpha)} \overset{\mathrm{d}}{=} \frac{T_{\alpha,k\alpha}}{B_{k\alpha,n-k\alpha}} = \frac{T_{\alpha,n}}{B_{k,n/\alpha-k}^{1/\alpha}}, \tag{2.2}$$

where the variables in each ratio are independent, and throughout $B_{a,b}$ denotes a $\mathrm{Beta}(a, b)$ random variable. Hence, the EPPF of the general case is given in (1.1) with $\Psi_{n,k}^{[\alpha]}$ expressed as $\Psi_{n,k}^{[\alpha]} = \mathbb{E}[h(T_\alpha) \mid K_n = k] = \mathbb{E}\big[h\big(Y_{\alpha,k\alpha}^{(n-k\alpha)}\big)\big]$. It is noteworthy that (2.2) indicates that Mittag-Leffler variables play a role in the general α class of Gibbs partitions. We can also use this to verify the recursion (see [15] and [21, Lemma 2.1])

$$\Psi_{n,k}^{[\alpha]} = \frac{k\alpha}{n} \Psi_{n+1,k+1}^{[\alpha]} + \left(1 - \frac{k\alpha}{n}\right) \Psi_{n+1,k}^{[\alpha]}.$$

Remarkably, for the number of blocks K_n having the probability mass function $\Psi_{n,k}^{[\alpha]} \mathbb{P}_{\alpha,0}^{(n)}(k)$, it follows from [32, Proposition 13] that as $n \to \infty$, $n^{-\alpha} K_n$ converges almost surely to $T^{-\alpha}$, which is the α-diversity under $\mathrm{PK}_\alpha(h \cdot f_\alpha)$ with density $h(s^{-1/\alpha}) g_\alpha(s)$, regardless of how complex $\Psi_{n,k}^{[\alpha]}$ may appear.

3. Fragmentation operator $\mathrm{FRAG}_{\alpha,-\beta}$

For $0 \leq \beta < \alpha$, we define the $\mathrm{PD}(\alpha, -\beta)$ fragmentation operator on $\mathcal{P}_\infty$ (and also Π_∞) as in [31], and more generally in [33, p. 108]. We refer to the operation on Π_∞ as the $\mathrm{CRP}(\alpha, -\beta)$ fragmentation operator. In both cases, we use $\mathrm{FRAG}_{\alpha,-\beta}$ to denote this operation in the respective contexts.

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

3.1. PD(α, $-\beta$) fragmentation on mass partitions

Let $(\hat{Q}_j^{(k)})_{j\geq 1} \sim \text{PD}(\alpha, -\beta))$ for $k=1, 2, \ldots$ denote a countably infinite collection of independent elements in $\mathcal{P}_\infty$, each having a PD(α, $-\beta$) distribution. Independent of this, select $\mathbf{V} := (V_k)_{k\geq 1} \in \mathcal{P}_\infty$. A PD($\alpha$, $-\beta$) fragmentation operation first shatters each mass V_k by an independent mass partition $(\hat{Q}_j^{(k)})_{j\geq 1} \sim \text{PD}(\alpha, -\beta)$ such that $V_k(\hat{Q}_j^{(k)})_{j\geq 1} = (V_k\hat{Q}_1^{(k)}, V_k\hat{Q}_2^{(k)}, \ldots)$ for $k=1, 2, \ldots$ The collection is then arranged in ranked order as an element in $\mathcal{P}_\infty$. The $\text{FRAG}_{\alpha,-\beta}$ operator is defined as follows, with Rank denoting the operation corresponding to the ranked rearrangement of masses:

$$\text{FRAG}_{\alpha,-\beta}(\mathbf{V}) = \text{Rank}\big(V_k\hat{Q}_j^{(k)}, j \geq 1, k \geq 1\big) \in \mathcal{P}_\infty. \tag{3.1}$$

3.2. CRP(α, $-\beta$) fragmentation on integer partitions

The corresponding (α, $-\beta$) fragmentation operation on the space of integer partitions Π_∞ for a given partition $\Pi_n = \{C_1, \ldots, C_{K_n}\}$ is generated from the corresponding $\mathbf{V}$. This operation shatters each block C_j by an independent CRP(α, $-\beta$) partition of the set C_j, with EPPF $p_{\alpha,-\beta}$ arising as a special case of $p_{\alpha,\theta}$ specified by (1.2). For given C_j, this results in a CRP(α, $-\beta$) partition of C_j, denoted as $\{\tilde{C}_{1,j}, \ldots, \tilde{C}_{K^{(j)}_{|C_j|},j}\}$, where the number of unique blocks $K^{(j)}_{|C_j|} \leq |C_j|$ corresponds to that of a CRP(α, $-\beta$) partition of C_j. The $\text{FRAG}_{\alpha,-\beta}$ operation on $\Pi_n = \{C_1, \ldots, C_{K_n}\}$ results in a partition of $[n]$ with $\sum_{j=1}^{K_n} K^{(j)}_{|C_j|} \leq n$ distinct blocks formed by the collection

$$\big\{\tilde{C}_{i,j}, i = 1, \ldots, K^{(j)}_{|C_j|}, j = 1, \ldots, K_n\big\}, \tag{3.2}$$

arranged according to the least element.

In the case where $\mathbf{V} \sim \text{PD}(\beta, \theta)$, [31] shows that $\tilde{\mathbf{V}} = \text{FRAG}_{\alpha,-\beta}(\mathbf{V}) \sim \text{PD}(\alpha, \theta)$. Conversely, as in [31, 33], applying an independent PD(β/α, θ/α) coagulation operation to the fragmented process $\tilde{\mathbf{V}}$ reverses this operation. The corresponding duality on integer partitions Π_∞ is described in more detail in [31, Theorem 12]. For a general treatment of these notions, see [3].

3.3. PD(α, $-\beta$) fragmentation of $\text{PK}_\beta(h \cdot f_\beta)$ mass partitions

We now address the question of the resulting distribution from the same independent fragmentation operation, $\text{FRAG}_{\alpha,-\beta}(\mathbf{V})$, when $\mathbf{V} := (V_k)_{k\geq 1} \sim \text{PK}_\beta(h \cdot f_\beta)$, where for clarity $h(t)$ is chosen with respect to T_β, having density f_β such that $\mathbb{E}[h(T_\beta)] = 1$. In order to achieve our results, we work with independent stable subordinators $\mathbf{T}_\alpha := (T_\alpha(t), t \geq 0)$ and $\mathbf{T}_{\beta/\alpha} := (T_{\beta/\alpha}(t), t \geq 0)$ generating independent elements of $\mathcal{P}_\infty$, $\mathbf{V}_\alpha \sim \text{PD}(\alpha, 0)$ and $\mathbf{Q}_{\beta/\alpha} \sim \text{PD}(\beta/\alpha, 0)$ respectively. The corresponding independent local-time processes starting at zero are $(L_\alpha(t), t \geq 0)$ and $(L_{\beta/\alpha}(t), t \geq 0)$, satisfying (2.1), with local times at 1, denoted respectively as $L_{1,\alpha} := L_\alpha(1) \overset{\text{d}}{=} T_\alpha^{-\alpha}$ and $L_{1,\beta/\alpha} := L_{\beta/\alpha}(1) \overset{\text{d}}{=} T_{\beta/\alpha}^{-\beta/\alpha}$, playing the role of L_1, as we have described, and otherwise following the more detailed description in [5] as it relates to the special case of the duality [31, Theorem 14 and Corollary 15]. The mass partition $\mathbf{V}_\beta \sim \text{PD}(\beta, 0)$ is formed by the independent coagulation $\mathbf{V}_\beta = \text{PD}(\beta/\alpha, 0) - \text{COAG}(\mathbf{V}_\alpha)$, which, as in [3–5] (for more details see the description in [33, Lemma 5.18, p. 114–115] in relation to the Bolthausen–Sznitman coalescent [6]), corresponds to the ranked normalized

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

jumps formed by the composition of the processes.

$$(F_\beta(u) = F_\alpha(F_{\beta/\alpha}(u)), u \in [0, 1]) \stackrel{d}{=} \left(\frac{T_\alpha(T_{\beta/\alpha}(u))}{T_\alpha(T_{\beta/\alpha}(1))}, u \in [0, 1]\right), \tag{3.3}$$

and has local time starting at 0 up to time 1 $L_{1,\beta} := L_\beta(1) = L_{\beta/\alpha}(L_{1,\alpha}) \stackrel{d}{=} T_\beta^{-\beta}$, with inverse local time up to time 1 $T_\beta := T_\beta(1) = T_\alpha(T_{\beta/\alpha}(1))$, with density f_β. Conversely, $\mathbf{V}_\alpha = \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}_\beta) \sim \mathrm{PD}(\alpha, 0)$.

Remark 3.1. We will write $F_\alpha \circ F_{\beta/\alpha}$ to indicate the composition of functions $F_\alpha(F_{\beta/\alpha}(u))$ for $u \in [0, 1]$.

Remark 3.2. There is the well-known distributional equivalence $T_\beta \stackrel{d}{=} T_\alpha \times T_{\beta/\alpha}^{1/\alpha} \stackrel{d}{=} T_{\beta/\alpha} \times T_\alpha^{\alpha/\beta}$. However, in the case of interpretation of the $\mathrm{PD}(\beta/\alpha, 0)$ coagulation, as in (3.3), the order matters and thus we will only use $T_\beta \stackrel{d}{=} T_\alpha \times T_{\beta/\alpha}^{1/\alpha}$.

Applying the scaling property in (2.1), it follows that, for any $s > 0$, $[L_{\beta/\alpha}(s^{-\alpha})]^{-1/\beta} \stackrel{d}{=} T_{\beta/\alpha}^{1/\alpha} s$, and also, using $L_{1,\alpha} \stackrel{d}{=} T_\alpha^{-\alpha}$ independent of the process $(L_{\beta/\alpha}(t), t \geq 0)$, we have

$$\omega_{\beta/\alpha,\beta}^{(y)}(s) := \frac{\mathbb{P}\left([L_{\beta/\alpha}(L_{1,\alpha})]^{-1/\beta} \in \mathrm{d}y \mid L_{1,\alpha}^{-1/\alpha} = s\right)}{\mathbb{P}\left(L_{1,\beta}^{-1/\beta} \in \mathrm{d}y\right)} = \frac{\alpha y^{\alpha-1} f_{\beta/\alpha}((y/s)^\alpha)}{s^\alpha f_\beta(y)} \tag{3.4}$$

such that the conditional distribution of $L_{1,\alpha} \mid L_{1,\beta}$ may be expressed in terms of that of the transformed variable $L_{1,\alpha}^{-1/\alpha} \mid L_{1,\beta}^{-1/\beta}$ as

$$\mathbb{P}\left(L_{1,\alpha}^{-1/\alpha} \in \mathrm{d}s \mid [L_{\beta/\alpha}(L_{1,\alpha})]^{-1/\beta} = y\right)/\mathrm{d}s = \omega_{\beta/\alpha,\beta}^{(y)}(s) f_\alpha(s), \tag{3.5}$$

which is equivalent to the conditional density of T_α given $T_\alpha \times T_{\beta/\alpha}^{1/\alpha} = y$. We now state our first result.

Theorem 3.1. *Let* $\mathbf{V} \sim \mathrm{PK}_\beta(h \cdot f_\beta)$ *with local time at 1, say* $L_{1,\mathbf{V}}$, *having density* $h(x^{-1/\beta}) g_\beta(x)$. *For any choice of* $0 < \beta < \alpha < 1$, *let* $\mathrm{FRAG}_{\alpha,-\beta}(\cdot)$ *denote a* $\mathrm{PD}(\alpha, -\beta)$ *fragmentation operator independent of* $\mathbf{V}$, *as defined in (3.1). Then:*

(*i*) $\tilde{\mathbf{V}} := \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}) \sim \mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$, *where*

$$\tilde{h}_{\beta/\alpha}(s) := \mathbb{E}_{\beta/\alpha}\left[h\left(s T_{\beta/\alpha}^{1/\alpha}\right)\right] = \int_0^\infty h(s y^{1/\alpha}) f_{\beta/\alpha}(y)\, \mathrm{d}y.$$

That is, it has a local time at time 1, say $L_{1,\tilde{\mathbf{V}}}$, *with density* $\tilde{h}_{\beta/\alpha}(z^{-1/\alpha}) g_\alpha(z)$.

(*ii*) *The conditional distribution of* $\tilde{\mathbf{V}} \mid L_{1,\mathbf{V}} = y^{-\beta}$ *is equivalent to the distribution of* $\mathbf{V}_\alpha \mid L_{1,\beta} = y^{-\beta}$, *where* $\mathbf{V}_\alpha = \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}_\beta) \sim \mathrm{PD}(\alpha, 0)$, *which is*

$$\mathrm{PD}_{\alpha|\beta}(\alpha \mid y) := \int_0^\infty \mathrm{PD}(\alpha \mid s) \omega_{\beta/\alpha,\beta}^{(y)}(s) f_\alpha(s)\, \mathrm{d}s = \mathrm{PK}_\alpha\left(\omega_{\beta/\alpha,\beta}^{(y)} \cdot f_\alpha\right) \tag{3.6}$$

for $\omega_{\beta/\alpha,\beta}^{(y)}(s)$ *defined in (3.4).*

(*iii*) *For clarity, when* $\mathbf{V}_\beta(y)$ *is equivalent in distribution to* $\mathbf{V}_\beta \mid T_\beta = y \sim \mathrm{PD}(\beta \mid y)$, *it follows that* $\mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}_\beta(y)) \sim \mathrm{PD}_{\alpha|\beta}(\alpha \mid y)$ *as in (3.6).*

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

Proof. Let $\hat{\mathbb{E}}^{(\beta,0)}_{(\alpha,-\beta)}$ denote the expectation with respect to the joint law of $(\mathbf{V}_\beta, ((\hat{Q}^{(k)})_{j\ge1}, k\ge1))$ where $\mathbf{V}_\beta \sim \mathrm{PD}(\beta, 0)$ with local time at 1 $L_{1,\beta} = L_{\beta/\alpha}(L_{1,\alpha})$, with density $g_\beta(x)$, and, independent of this, $((\hat{Q}^{(k)})_{j\ge1}, k\ge1)$ are iid $\mathrm{PD}(\alpha, -\beta)$ mass partitions. Consider $\mathbf{V} \sim \mathrm{PK}_\beta(h\cdot f_\beta)$. The distribution of $\tilde{\mathbf{V}} = \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V})$ is characterized, for a measurable function Φ, by

$$\mathbb{E}[\Phi(\mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}))] = \hat{\mathbb{E}}^{(\beta,0)}_{(\alpha,-\beta)}\Big[\Phi(\mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}_\beta))h\big(L_{1,\beta}^{-1/\beta}\big)\Big]. \tag{3.7}$$

But, from [5, 31], for $\mathbf{V}_\alpha \sim \mathrm{PD}(\alpha, 0)$ this is equivalent to

$$\mathbb{E}\Big[\Phi(\mathbf{V}_\alpha)h([L_{\beta/\alpha}(L_{1,\alpha})]^{-1/\beta})\Big]. \tag{3.8}$$

Now use (3.4) and (3.5) to write the joint density of $L_{1,\alpha}^{-1/\alpha}$ and $[L_{\beta/\alpha}(L_{1,\alpha})]^{-1/\beta}$ for the respective arguments (s,y) as $\omega^{(y)}_{\beta/\alpha,\beta}(s)f_\alpha(s)f_\beta(y)$. Then, conditioning on the joint event $L_{1,\alpha}^{-1/\alpha} = s$ and $[L_{\beta/\alpha}(L_{1,\alpha})]^{-1/\beta} = y$, we can express (3.8) as

$$\int_0^\infty\int_0^\infty \mathbb{E}\big[\Phi(\mathbf{V}_\alpha)\mid L_{1,\alpha}^{-1/\alpha} = s,\ [L_{\beta/\alpha}(L_{1,\alpha})]^{-1/\beta} = y\big]h(y)\omega^{(y)}_{\beta/\alpha,\beta}(s)f_\alpha(s)f_\beta(y)\,\mathrm{d}s\,\mathrm{d}y. \tag{3.9}$$

It follows that, since $(L_{\beta/\alpha}(t), t\ge0)$ is independent of $\mathbf{V}_\alpha$, $\mathbf{V}_\alpha \mid L_{1,\alpha}^{-1/\alpha} = s$, $[L_{\beta/\alpha}(L_{1,\alpha})]^{-1/\beta} = y$ has distribution $\mathrm{PD}(\alpha\mid s)$. Using this and (3.9), it follows that the expectation in (3.8) can be expressed as

$$\int_0^\infty\bigg[\int_0^\infty \mathbb{E}\big[\Phi(\mathbf{V}_\alpha)\mid T_\alpha = s\big]\omega^{(y)}_{\beta/\alpha,\beta}(s)f_\alpha(s)\,\mathrm{d}s\bigg]h(y)f_\beta(y)\,\mathrm{d}y. \tag{3.10}$$

Alternatively, it can also be expressed as $\int_0^\infty \mathbb{E}[\Phi(\mathbf{V}_\alpha)\mid T_\alpha = s]\tilde{h}_{\beta/\alpha}(s)f_\alpha(s)\,\mathrm{d}s$ for $\tilde{h}_{\beta/\alpha}(s) := \mathbb{E}_{\beta/\alpha}\big[h\big(sT^{1/\alpha}_{\beta/\alpha}\big)\big] = \int_0^\infty \omega^{(y)}_{\beta/\alpha,\beta}(s)h(y)f_\beta(y)\,\mathrm{d}y$, yielding the results in (i) and (ii). Now comparing $\hat{\mathbb{E}}^{(\beta,0)}_{(\alpha,-\beta)}\big[\Phi(\mathrm{FRAG}_{\alpha,-\beta}(V_\beta))h\big(L_{1,\beta}^{-1/\beta}\big)\big]$ in (3.7) with the equivalent expression in (3.10), it follows that, for almost all y,

$$\hat{\mathbb{E}}^{(\beta,0)}_{(\alpha,-\beta)}\big[\Phi(\mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}_\beta))\mid T_\beta = y\big] = \int_0^\infty \mathbb{E}[\Phi(\mathbf{V}_\alpha)\mid T_\alpha = s]\omega^{(y)}_{\beta/\alpha,\beta}(s)f_\alpha(s)\,\mathrm{d}s,$$

and, noting the independence of the $\mathrm{FRAG}_{\alpha,-\beta}$ operator relative to $\mathbf{V}_\beta$, this leads to the result in (iii). □

4. Duality via dependent coagulation

We now describe how to construct dependent coagulations to complete the dual process of recovering $\mathbf{V} \sim \mathrm{PK}_\beta(h\cdot f_\beta)$ from the coagulation of $\tilde{\mathbf{V}} = \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}) \sim \mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha}\cdot f_\alpha)$ by some mass partition $\mathbf{Q} := (Q_\ell)_{\ell\ge1} \in \mathcal{P}_\infty$. Our results show how specification of $h(t)$ leads to a prescription for identifying the laws of $\mathbf{V}$, $\tilde{\mathbf{V}}$, and $\mathbf{Q}$ without guesswork.

Recall that for the independent mass partitions $(\mathbf{V}_\alpha, \mathbf{Q}_{\beta/\alpha})$ described in (3.3), the process of coagulation yields an inverse local time at 1 for $\mathbf{V}_\beta$ to be $T_\beta(1) = T_\alpha(T_{\beta/\alpha}(1)) \overset{\mathrm{d}}{=} T_\alpha \times T^{1/\alpha}_{\beta/\alpha}$. For $\tilde{\mathbf{V}}$ as described above, we consider the dependent pair $(\tilde{\mathbf{V}}, \mathbf{Q})$ with joint law, say $\mathrm{P}^{\beta/\alpha}_\alpha(h)$, characterized by, for some generic measurable function Φ,

$$\mathbb{E}[\Phi(\tilde{\mathbf{V}}, \mathbf{Q})] = \mathbb{E}^{(\beta/\alpha,0)}_{(\alpha,0)}\big[\Phi(\mathbf{V}_\alpha, \mathbf{Q}_{\beta/\alpha})h(T_\alpha(T_{\beta/\alpha}(1)))\big], \tag{4.1}$$

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

with $\mathbb{E}[h(T_\beta(1))] = \mathbb{E}[h(T_\alpha(T_{\beta/\alpha}(1)))] = 1$, and the notation $\mathbb{E}^{(\beta/\alpha,0)}_{(\alpha,0)}$ referring to an expectation evaluated under the joint law of the independent PD(α, 0) and PD(β/α, 0) distributions. We use this for clarity, but will suppress it when it is clear we are referring to such variables. Equivalently, by conditioning and scaling properties, the joint law of $(\tilde{\mathbf{V}}, \mathbf{Q})$ is given by

$$\mathrm{P}^{\beta/\alpha}_{\alpha}(h) := \int_0^\infty \int_0^\infty \mathrm{PD}(\alpha \mid s)\mathrm{PD}(\beta/\alpha \mid y) h(sy^{1/\alpha}) f_{\beta/\alpha}(y) f_\alpha(s)\, \mathrm{d}y\, \mathrm{d}s. \tag{4.2}$$

Remark 4.1. Throughout, let $(T_{\tilde{\mathbf{V}}}, T_{\mathbf{Q}})$ denote the generalization of $(T_\alpha, T_{\beta/\alpha})$ for $(\tilde{\mathbf{V}}, \mathbf{Q})$. Equation (4.2) indicates that their joint density is given by $h(sy^{1/\alpha}) f_{\beta/\alpha}(y) f_\alpha(s)$ and determines the overall dependency of the processes.

In addition, for collections of iid Uniform[0, 1] variables $((\tilde{U}_k), (U_\ell))$ independent of $(\tilde{\mathbf{V}}, \mathbf{Q})$, define random distribution functions (exchangeable bridges), for $y \in [0, 1]$,

$$F_{\tilde{\mathbf{V}}}(y) = \sum_{k=1}^{\infty} \tilde{V}_k \mathbb{1}_{\{\tilde{U}_k \le y\}}, \qquad F_{\mathbf{Q}}(y) = \sum_{\ell=1}^{\infty} Q_\ell \mathbb{1}_{\{U_\ell \le y\}}. \tag{4.3}$$

Remark 4.2. It follows that when $h(t) = t^{-\theta}/\mathbb{E}\big[T_\beta^{-\theta}\big]$ for $\theta > -\beta$, $\tilde{\mathbf{V}} \sim \mathrm{PD}(\alpha, \theta)$ is independent of $\mathbf{Q} \sim \mathrm{PD}(\beta/\alpha, \theta/\alpha)$. Hence, $F_{\tilde{\mathbf{V}}} \overset{\mathrm{d}}{=} F_{\alpha,\theta}$ and $F_{\mathbf{Q}} \overset{\mathrm{d}}{=} F_{\beta/\alpha,\theta/\alpha}$ are independent.

Proposition 4.1. *For $0 < \beta < \alpha < 1$, let $(\tilde{\mathbf{V}}, \mathbf{Q})$ have a joint distribution $\mathrm{P}^{\beta/\alpha}_{\alpha}(h)$ specified by (4.1), or equivalently (4.2), such that $\tilde{\mathbf{V}} \sim \mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$ and $(F_{\tilde{\mathbf{V}}}, F_{\mathbf{Q}})$ are bridges defined in (4.3), and $(T_{\tilde{\mathbf{V}}}, T_{\mathbf{Q}})$ as in Remark 4.1. Let $\mathbf{V} \in \mathcal{P}_\infty$ be the ranked masses of the bridge defined by the composition $F_{\mathbf{V}} := F_{\tilde{\mathbf{V}}} \circ F_{\mathbf{Q}}$. Then, $\mathbf{V}$ is equivalent to the coagulation of $\tilde{\mathbf{V}}$ by $\mathbf{Q}$ and there are the following properties:*

(i) $\mathbf{V} \sim \mathrm{PK}_\beta(h \cdot f_\beta)$.

(ii) *The marginal distribution of $\mathbf{Q} \sim \mathrm{PK}_{\beta/\alpha}(\hat{h}_\alpha \cdot f_{\beta/\alpha})$, where*

$$\hat{h}_\alpha(y) := \mathbb{E}_\alpha[h(T_\alpha y^{1/\alpha})] = \int_0^\infty h(sy^{1/\alpha}) f_\alpha(s)\, \mathrm{d}s,$$

and the corresponding $T_{\mathbf{Q}}$ has density $\hat{h}_\alpha(y) f_{\beta/\alpha}(y)$.

(iii) *The distribution of $\tilde{\mathbf{V}} \mid T_{\mathbf{Q}} = y$ is $\mathrm{PK}_\alpha(h^{(y)}_\alpha \cdot f_\alpha)$, where $h^{(y)}_\alpha(s) = h(sy^{1/\alpha})/\mathbb{E}_\alpha[h(T_\alpha y^{1/\alpha})]$.*

Proof. We first recall from (3.3) that under independent PD(α, 0) and PD(β/α, 0) laws, the bridge $F_\beta := F_\alpha \circ F_{\beta/\alpha}$ follows the law of a PD(β, 0)-bridge with inverse local time at 1, $T_\beta := T_\beta(1) = T_\alpha(T_{\beta/\alpha}(1))$. Hence, under the joint law of $(\tilde{\mathbf{V}}, \mathbf{Q})$ specified by (4.1), it follows that, for $F_{\mathbf{V}} := F_{\tilde{\mathbf{V}}} \circ F_{\mathbf{Q}}$,

$$\mathbb{E}[\Phi(F_{\tilde{\mathbf{V}}} \circ F_{\mathbf{Q}})] = \mathbb{E}\big[\Phi(F_\alpha \circ F_{\beta/\alpha}) h(T_\alpha(T_{\beta/\alpha}(1)))\big] = \mathbb{E}[\Phi(F_\beta) h(T_\beta(1))],$$

showing that $F_{\mathbf{V}}$ is a $\mathrm{PK}_\beta(h \cdot f_\beta)$-bridge and thus $\mathbf{V} \sim \mathrm{PK}_\beta(h \cdot f_\beta)$ in statement (i). Statements (ii) and (iii) follow from straightforward usage of (4.2). ☐

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

5. Gibbs partitions of $[n]$ derived from $\mathrm{FRAG}_{\alpha,-\beta}$

Recall from [32, 33] that when $\mathbf{V}_\beta \sim \mathrm{PD}(\beta, 0)$, $\mathbf{V}_\beta \mid L_{1,\beta} = y^{-\beta}$ is equivalent in distribution to $\mathbf{V}_\beta \mid T_\beta = y \sim \mathrm{PD}(\beta \mid y)$ and has the associated Gibbs partition of $[n]$ described by the $\mathrm{PD}(\beta \mid y)$-EPPF

$$p_\beta(n_1, \ldots, n_k \mid y) := \frac{f^{(n-k\beta)}_{\beta,k\beta}(y)}{f_\beta(y)} p_\beta(n_1, \ldots, n_k), \tag{5.1}$$

where, as in [20, 21],

$$\frac{f^{(n-k\beta)}_{\beta,k\beta}(y)}{f_\beta(y)} = \mathbb{G}^{(n,k)}_\beta(y) \frac{\beta^{1-k}\Gamma(n)}{\Gamma(k)},$$

with, from [15, 32, 33],

$$\mathbb{G}^{(n,k)}_\beta(t) = \frac{\beta^k t^{-n}}{\Gamma(n-k\beta) f_\beta(t)} \left[\int_0^t f_\beta(v)(t-v)^{n-k\beta-1}\, \mathrm{d}v \right], \tag{5.2}$$

and $f^{(n-k\beta)}_{\beta,k\beta}(y)$ being the conditional density of $T_\beta \mid K^{[\beta]}_n = k$ corresponding to a random variable denoted as $Y^{n-k\beta}_{\beta,k\beta}$, as otherwise described in (2.2) with β in place of α. Note, furthermore, as in [21], this means that $T_\beta := T_\alpha(T_{\beta/\alpha}(1)) \overset{\mathrm{d}}{=} Y^{(n-K^{[\beta]}_n\beta)}_{\beta,K^{[\beta]}_n\beta}$ for $K^{[\beta]}_n \sim \mathbb{P}^{(n)}_{\beta,0}(k)$. We use these facts to obtain interesting expressions for α-Gibbs partitions equivalent to those arising from the $\mathrm{FRAG}_{\alpha,-\beta}$ operator. In particular, in Section 5.1 we provide a remarkable description of the corresponding EPPF for mass partitions equivalent in distribution to

$$\mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}_\beta(y)) \sim \mathrm{PD}_{\alpha|\beta}(\alpha \mid y) = \mathrm{PK}_\alpha\big(\omega^{(y)}_{\beta/\alpha,\beta} \cdot f_\alpha\big) \tag{5.3}$$

for $\mathbf{V}_\beta(y) \sim \mathrm{PD}(\beta \mid y)$, as described in the result in Theorem 3.1(iii).

5.1. Gibbs partitions of $[n]$ of $V_\alpha \mid L_{1,\beta}$, equivalently of $\mathrm{FRAG}_{\alpha,-\beta}(V) \mid L_{1,V}$

Recall from Theorem 3.1 that the distribution of $\mathbf{V}_\alpha \mid L_{1,\beta} = y^{-\beta}$ is equivalent to that of $\tilde{\mathbf{V}} = \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}) \mid L_{1,V} = y^{-\beta}$, with distribution denoted $\mathrm{PD}_{\alpha|\beta}(\alpha \mid y) := \mathrm{PK}_\alpha\big(\omega^{(y)}_{\beta/\alpha,\beta} \cdot f_\alpha\big)$ as in (3.6), where $\omega^{(y)}_{\beta/\alpha,\beta}(s)$ is a ratio of stable densities and hence does not have an explicit form for general $0 < \beta < \alpha < 1$. We now present results for the EPPF of the $\mathrm{PD}_{\alpha|\beta}(\alpha \mid y)$ Gibbs partition of $[n]$. We first note that since $T_\alpha \mid K^{[\alpha]}_n = k$ is equivalent in distribution to $Y^{n-k\alpha}_{\alpha,k\alpha}$ with density $f^{(n-k\alpha)}_{\alpha,k\alpha}$, the EPPF can be expressed as $\big[\int_0^\infty \omega^{(y)}_{\beta/\alpha,\beta}(s) f^{n-k\alpha}_{\alpha,k\alpha}(s)\, \mathrm{d}s\big] p_\alpha(n_1, \ldots, n_k)$, where the first integral term is the density of $Y^{(n-k\alpha)}_{\alpha,k\alpha} \times T^{1/\alpha}_{\beta/\alpha}$ divided by $f_\beta(y)$, and does not have an obvious recognizable form. However, we can use the approach in [20] to express $\omega^{(y)}_{\beta/\alpha,\beta}$ in terms of Fox-H functions [29], leading to the expression for the EPPF in terms of Fox-H functions given in Appendix 10.

The next result provides a more revealing expression that is not obvious.

Theorem 5.1. *The EPPF of the* $\mathrm{PD}_{\alpha|\beta}(\alpha \mid y)$ *Gibbs partition of* $[n]$*, corresponding to sampling* n *times from mass partitions with distribution described in (5.3), can be expressed as*

$$p_{\alpha|\beta}(n_1, \ldots, n_k \mid y) := \left[\sum_{j=1}^k \mathbb{P}^{(k)}_{\beta/\alpha,0}(j) \frac{f^{(n-j\beta)}_{\beta,j\beta}(y)}{f_\beta(y)} \right] p_\alpha(n_1, \ldots, n_k), \tag{5.4}$$

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

where $\mathbb{P}^{(k)}_{\beta/\alpha,0}(j) = \mathbb{P}_{\beta/\alpha,0}(K_k = j)$ *is the distribution of the number of blocks in a* PD$(\beta/\alpha, 0)$ *partition of [k], and* $\sum_{j=1}^{k} \mathbb{P}^{(k)}_{\beta/\alpha,0}(j) f^{(n-j\beta)}_{\beta,j\beta}(y)$ *is the conditional density of* $T_\beta \mid K^{[\alpha]}_n = k$ *for the number of blocks* $K^{[\alpha]}_n$ *in a* PD$(\alpha, 0)$ *partition of [n], with* $T_\beta := T_\alpha(T_{\beta/\alpha}(1)) \overset{\mathrm{d}}{=} T_\alpha \times T^{1/\alpha}_{\beta/\alpha}$ *being equivalent to the inverse local time at 1 of* $\mathbf{V}_\beta \sim$ PD$(\beta, 0)$.

Proof. The expression for the EPPF is the conditional distribution of a PD$(\alpha, 0)$ partition of $[n]$ given $T_\beta = y$. The joint distribution may be expressed as in (5.4) in terms of the marginal EPPF $p_\alpha(n_1, \ldots, n_k)$ and the conditional density of $T_\beta \mid K^{[\alpha]}_n = k$. It remains to show that $T_\beta \mid K^{[\alpha]}_n = k$ agrees with the expression in (5.4) as indicated. Recall that $L_{1,\beta} = L_{\beta/\alpha}(L_{1,\alpha})$ and hence the corresponding inverse local time at 1 is $T_\beta := T_\beta(1) = T_\alpha(T_{\beta/\alpha}(1))$, corresponding to the coagulation operation dictated by $F_\beta = F_\alpha \circ F_{\beta/\alpha}$, as expressed in (3.3). Sampling from $F_\alpha \circ F_{\beta/\alpha}$, that is, according to variables $\big(F^{-1}_{\beta/\alpha}(F^{-1}_\alpha(U'_i)), i \in [n]\big)$, where F^{-1} denotes the quantile function, it follows that this procedure produces a PD$(\beta, 0)$ partition of $[n]$ with $K^{[\beta]}_n \overset{\mathrm{d}}{=} K^{[\beta/\alpha]}_{K^{[\alpha]}_n}$ blocks, where the two components are independent. Furthermore, the order matters, giving $K^{[\alpha]}_n$ the interpretation as the number of blocks to be merged, according to a PD$(\beta/\alpha, 0)$ partition of $[k]$, for $K^{[\alpha]}_n = k \le n$. Now, from [21], $T_\beta \overset{\mathrm{d}}{=} Y^{n-K^{[\beta]}_n\beta}_{\beta,K^{[\beta]}_n\beta}$. Hence $T_\beta \mid K^{[\alpha]}_n = k$ is equivalent to $Y^{(n-K^{[\beta/\alpha]}_k\beta)}_{\beta,K^{[\beta/\alpha]}_k\beta}$, which, using (5.1), leads to the description of the density of $T_\beta \mid K^{[\alpha]}_n = k$ appearing in (5.4). □

Remark 5.1. The previous result is equivalent to showing that $Y^{(n-K^{[\beta/\alpha]}_k\beta)}_{\beta,K^{[\beta/\alpha]}_k\beta} \overset{\mathrm{d}}{=} Y^{(n-k\alpha)}_{\alpha,k\alpha)} \times T^{1/\alpha}_{\beta/\alpha}$, which can be deduced directly using the subordinator representation [21, Theorem 2.1 and Proposition 2.1] and decompositions of beta variables.

5.2. EPPF of $\mathrm{FRAG}_{\alpha,-\beta}(V) \sim \mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$

Recall from [15, 32] (see also [21]) that if $\mathbf{V} \sim \mathrm{PK}_\beta(h \cdot f_\beta)$ with $\mathbb{E}[h(T_\beta)] = 1$, then the EPPF of its associated Gibbs partition of $[n]$ is described as $p^{[\gamma]}_\beta(n_1, \ldots, n_k) = \Psi^{[\beta]}_{n,k} \times p_\beta(n_1, \ldots, n_k)$, where $\Psi^{[\beta]}_{n,k} = \mathbb{E}_\beta\big[h(T_\beta) \mid K^{[\beta]}_n = k\big]$ and, for clarity, $K^{[\beta]}_n$ is the number of blocks of a PD$(\beta, 0)$ partition of $[n]$.

Theorem 5.1 leads to the EPPF corresponding to $\tilde{\mathbf{V}} = \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}) \sim \mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$, or any variable in $\mathcal{P}_\infty$ having the same distribution.

Proposition 5.1. *Suppose that, for* $0 < \beta < \alpha < 1$, $\tilde{\mathbf{V}} \sim \mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$, *where* $\tilde{h}_{\beta/\alpha}(v) := \mathbb{E}_{\beta/\alpha}\big[h\big(vT^{1/\alpha}_{\beta/\alpha}\big)\big]$. *Then, the* $\mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$ *EPPF of the associated Gibbs partition of [n] can be expressed as*

$$\left[\sum_{j=1}^{k} \mathbb{P}^{(k)}_{\beta/\alpha,0}(j) \Psi^{[\beta]}_{n,j}\right] p_\alpha(n_1, \ldots, n_k), \tag{5.5}$$

and we have the identity, for $T_\beta := T_\beta(1) = T_\alpha(T_{\beta/\alpha}(1))$,

$$\mathbb{E}_\alpha\big[\tilde{h}_{\beta/\alpha}(T_\alpha) \mid K^{[\alpha]}_n = k\big] = \mathbb{E}\big[h(T_\beta) \mid K^{[\alpha]}_n = k\big] = \sum_{j=1}^{k} \mathbb{P}^{(k)}_{\beta/\alpha,0}(j) \Psi^{[\beta]}_{n,j}.$$

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

Proof. The EPPF is equivalent to $\int_0^\infty p_{\alpha|\beta}(n_1, \ldots, n_k \mid y)h(y)f_\beta(y)\,dy$, and hence the result follows from (5.4) in Theorem 5.1. □

Remark 5.2. The expression in (5.5) provides a description of any mass partition with distribution $\mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$ where $\tilde{h}_{\beta/\alpha}(v) := \mathbb{E}_{\beta/\alpha}\big[h\big(vT_{\beta/\alpha}^{1/\alpha}\big)\big]$, regardless of whether or not it actually arises from a fragmentation operation.

As a check, in the case where $\tilde{\mathbf{V}} \overset{d}{=} \mathbf{V}_\alpha \sim \mathrm{PD}(\alpha, \theta)$ and hence $\mathbf{V} \overset{d}{=} \mathbf{V}_\beta \sim \mathrm{PD}(\beta, \theta)$, (5.5) must satisfy

$$\sum_{j=1}^{k} \mathbb{P}_{\beta/\alpha,0}^{(k)}(j) \frac{\Gamma(\theta/\beta + j)}{\Gamma(\theta/\beta + 1)\Gamma(j)} = \frac{\Gamma(\theta/\alpha + k)}{\Gamma(\theta/\alpha + 1)\Gamma(k)}.$$

However, this *is* satisfied since it agrees with [33, Exercise 3.2.9, p. 66], with k in place of n. We have the following corollary in the case of $\beta/\alpha = \frac{1}{2}$.

Corollary 5.1. *Specializing Theorem 5.1 to the case of* $\beta/\alpha = \frac{1}{2}$*, where* $\mathbf{V} \sim \mathrm{PK}_{\alpha/2}(h \cdot f_{\alpha/2})$ *and* $\Psi_{n,j}^{[\alpha/2]} = \mathbb{E}_{\alpha/2}\big[h(T_{\alpha/2}) \mid K_n^{[\alpha/2]} = j\big]$*, the* $\mathrm{PK}_\alpha(\tilde{h}_{1/2} \cdot f_\alpha)$ *EPPF in (5.5) becomes*

$$\left[\sum_{j=1}^{k} \binom{2k-j-1}{k-1} 2^{j+1-2k} \Psi_{n,j}^{[\alpha/2]}\right] p_\alpha(n_1, \ldots, n_k).$$

5.3. $\mathrm{PD}\big(\alpha, -\frac{1}{2}\big)$ fragmentation of a Brownian excursion partition conditioned on its local time

Following [32, Section 8] and [33, Section 4.5, p. 90], let $(P_{\ell,0})_{\ell \geq 1} \sim \mathrm{PD}\big(\frac{1}{2}, 0\big)$ denote the ranked excursion lengths of a standard Brownian motion $B := (B_t \colon t \in [0, 1])$, with corresponding local time at 0 up to time 1 given by $\tilde{L}_1 \overset{d}{=} (2T_{1/2})^{-1/2} \overset{d}{=} |B_1|$. Then, it follows that $(P_{\ell,0})_{\ell \geq 1} \mid \tilde{L}_1 = s$ has a $\mathrm{PD}\big(\frac{1}{2} \mid \frac{1}{2}s^{-2}\big)$ distribution. Furthermore, with respect to $(P_\ell(s))_{\ell \geq 1} \sim \mathrm{PD}\big(\frac{1}{2} \mid \frac{1}{2}s^{-2}\big)$, we describe the special $\beta = \frac{1}{2}$ explicit case of the Gibbs partitions (EPPF) of $[n]$ in terms of Hermite functions as derived in [32] (see also [33, Section 4.5]) as

$$p_{1/2}\big(n_1, \ldots, n_k \mid \tfrac{1}{2}s^{-2}\big) = s^{k-1} \tilde{H}_{k+1-2n}(s) \frac{\Gamma(n)}{2^{1-n}\Gamma(k)} p_{\frac{1}{2}}(n_1, \ldots, n_k), \tag{5.6}$$

where, for $U(a,b,c)$ a confluent hypergeometric function of the second kind (see [27, p. 263]),

$$\tilde{H}_{-2q}(s) = 2^{-q} U\left(q, \frac{1}{2}, \frac{s^2}{2}\right) = \sum_{\ell=0}^{\infty} \frac{(-s)^\ell}{\ell!} \frac{\Gamma(q + \ell/2)}{2\Gamma(2q)} 2^{q+\ell/2}$$

is a Hermite function of index $-2q$. That is, in (5.2), $\mathbb{G}_{1/2}^{(n,k)}\big(\frac{1}{2}s^{-2}\big) = 2^{n-k} s^{k-1} \tilde{H}_{k+1-2n}(s)$.

Proposition 5.2. *Suppose that* $\mathbf{P}_{1/2}(s) := (P_\ell(s))_{\ell \geq 1} \sim \mathrm{PD}\big(\frac{1}{2} \mid \frac{1}{2}s^{-2}\big)$*. Then, for* $\alpha > \frac{1}{2}$*,* $\mathbf{P}_{\alpha|1/2}(s) := \mathrm{FRAG}_{\alpha,-1/2}(\mathbf{P}_{1/2}(s)) \sim \mathrm{PD}_{\alpha|1/2}\big(\alpha \mid \frac{1}{2}s^{-2}\big)$*, with corresponding EPPF expressed in terms of a mixture of Hermite functions,*

$$\left[\sum_{j=1}^{k} \mathbb{P}_{1/2\alpha,0}^{(k)}(j) 2^{n-1} s^{j-1} \tilde{H}_{j+1-2n}(s) \frac{\Gamma(n)}{\Gamma(j)}\right] p_\alpha(n_1, \ldots, n_k). \tag{5.7}$$

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

Proof. The result follows as a special case of Theorems 3.1 and 5.1, and otherwise using the explicit form of the EPPF in (5.6). □

Remark 5.3. In order to obtain a partition of $[n]$ corresponding to the EPPF in (5.7), we can sample from (5.6) via the prediction rules indicated in [32, (111) and (112)], and then employ the $\mathrm{CRP}\big(\alpha, -\frac{1}{2}\big)$ fragmentation scheme in (3.2).

5.4. $\mathrm{FRAG}_{\alpha,-\beta}$ for the Mittag-Leffler class

We now present results for an application of the $\mathrm{FRAG}_{\alpha,-\beta}$ operator to the most basic case of the Mittag-Leffler class as described in [21] (see also [23]). Consider again $\mathbf{V}_\beta \sim \mathrm{PD}(\beta, 0)$. Recall that for $\lambda > 0$, the Laplace transform of $L_{1,\beta} \overset{\mathrm{d}}{=} T_\beta^{-\beta}$ equates to the Mittag-Leffler function (see, for instance, [16]), expressed as

$$\mathrm{E}_{\beta,1}(-\lambda) = \mathbb{E}\big[\mathrm{e}^{-\lambda T_\beta^{-\beta}}\big] = \sum_{\ell=0}^{\infty} \frac{(-\lambda)^\ell}{\Gamma(\beta\ell+1)}.$$

Independent of $\mathbf{V}_\beta$, let $(N(s), s \geq 0)$ denote a standard Poisson process, where $\mathbb{E}[N(s)] = s$, and consider the mixed Poisson process $(N(tL_{1,\beta}), t \geq 0)$. Then, as shown in [21], $\mathbf{V}_\beta \mid N(\lambda L_{1,\beta}) = 0$ has a $\mathrm{PK}_\beta(h \cdot f_\beta)$ distribution of the form

$$\int_0^\infty \mathrm{PD}(\beta \mid t) \frac{\mathrm{e}^{-\lambda t^{-\beta}}}{\mathrm{E}_{\beta,1}(-\lambda)} f_\beta(t)\, \mathrm{d}t, \tag{5.8}$$

where $h(t) = \mathrm{e}^{-\lambda t^{-\beta}} / \mathrm{E}_{\beta,1}(-\lambda)$. Furthermore, from [21, Proposition 4.5] its corresponding EPPF for a partition of $[n]$ can be expressed as

$$\frac{\mathrm{E}_{\beta,n}^{(k)}(-\lambda)}{\mathrm{E}_{\beta,1}(-\lambda)} p_\beta(n_1, \ldots, n_k),$$

where, as in [21, Proposition 4.2],

$$\mathrm{E}_{\beta,n}^{(k)}(-\lambda) = \sum_{\ell=0}^{\infty} \frac{(-\lambda)^\ell}{\ell!} \frac{\Gamma(k+\ell)\Gamma(n)}{\Gamma(k)\Gamma(\beta\ell+n)}.$$

We now describe properties of the fragmented process; additional details to verify the calculations are provided within the statement of the result.

Proposition 5.3. *Let* $\mathbf{V}$ *have distribution specified in (5.8), and otherwise consider the setting in Theorem 3.1.*

(i) $\tilde{\mathbf{V}} = \mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V}) \sim \mathrm{PK}_\alpha(\tilde{h}_{\beta/\alpha} \cdot f_\alpha)$*, where*

$$\tilde{h}_{\beta/\alpha}(s) = \frac{\mathrm{E}_{\beta/\alpha,1}(-\lambda s^{-\beta})}{\mathrm{E}_{\beta,1}(-\lambda)} = \frac{\int_0^\infty \mathrm{e}^{-\lambda s^{-\beta} y^{-\beta/\alpha}} f_{\beta/\alpha}(y)\, \mathrm{d}y}{\mathrm{E}_{\beta,1}(-\lambda)}.$$

(ii) *Its EPPF for a partition of* $[n]$ *can be expressed as*

$$\frac{\sum_{\ell=0}^{\infty} \frac{(-\lambda)^\ell}{\Gamma(\beta\ell/\alpha+1)} \frac{\Gamma(n)\Gamma(\beta\ell/\alpha+k)}{\Gamma(k)\Gamma(\beta\ell+n)}}{\mathrm{E}_{\beta,1}(-\lambda)} p_\alpha(n_1, \ldots, n_k),$$

which follows from the use of $\mathbb{E}[T_\alpha^{-\beta\ell} \mid K_n = k]$ *in (1.2).*

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

(iii) *From (5.5) in Proposition 5.1, we have the identity*

$$\sum_{j=1}^{k}\mathbb{P}^{(k)}_{\beta/\alpha,0}(j)\mathrm{E}^{(j)}_{\beta,n}(-\lambda)=\sum_{\ell=0}^{\infty}\frac{(-\lambda)^{\ell}}{\Gamma(\beta\ell/\alpha+1)}\frac{\Gamma(n)\Gamma(\beta\ell/\alpha+k)}{\Gamma(k)\Gamma(\beta\ell+n)}.$$

6. Coagulation and fragmentation of generalized gamma models

For any $0<\beta<1$, let $(\tau_\beta(s),\, s\geq 0)$ denote a generalized gamma subordinator specified by its Laplace transform: $\mathbb{E}[\mathrm{e}^{-w\tau_\beta(s)}]=\mathrm{e}^{-s[(1+w)^{\beta}-1]}$. The generalized gamma subordinator, along with corresponding mass partitions and bridges defined by normalization, say $(\tau_\beta(\zeta u)/\tau_\beta(\zeta)\colon u\in[0,1])$ for $\zeta>0$, as described in [32], arises in numerous contexts. However, for the purpose of this exposition, the reader may refer to its role in the construction of $\mathrm{PD}(\beta,\theta)$ distributions for $\theta>0$, as described in [37, Proposition 21].

Based on notes provided by Jim Pitman [34], with relevance to species sampling and latent feature models, we can relate this and a more general size-biased class to $\mathbf{V}_\beta\sim\mathrm{PD}(\beta,0)$ as follows. Again, let $(N(s)\colon s\geq 0)$ denote a standard Poisson process independent of $\mathbf{V}_\beta$. Then the distribution of $\mathbf{V}_\beta\mid N(\zeta^{1/\beta}T_\beta(1))=m$ corresponds to the size-biased general gamma $\mathbf{V}:=(V_k)_{k\geq 1}\sim\mathbb{P}^{[m]}_\beta(\zeta):=\mathrm{PK}_\beta\big(r^{[m]}_{\beta,\zeta}\cdot f_\beta\big)$, where $r^{[m]}_{\beta,\zeta}(t)=t^m\mathrm{e}^{-\zeta^{1/\beta}t}/\mathbb{E}\big[T^m_\beta\mathrm{e}^{-\zeta^{1/\beta}T_\beta}\big]$ for $m=0,1,2,\ldots$, as described in [21, 24, 25].

Remark 6.1. Note that, from [34], within a species sampling context, $N(\zeta^{1/\beta}T_\beta(1))$ can be interpreted as the number of animals trapped/sampled according to the process of sampling from $F_{\mathbf{V}_\beta}$ up to time $\zeta^{1/\beta}$.

The case of $m=0$ corresponds to the well-known case of the distribution of the mass partitions of $(\tau_\beta(\zeta u)/\tau_\beta(\zeta)\colon u\in[0,1])$. The other cases are less studied but arise in various contexts. Here, for brevity, we show how to use Proposition 4.1 to easily identify laws and explicit constructions of (dependent) random measures leading to a coagulation/fragmentation duality in the case of $m=1$. That is, suppose that $\mathbf{V}\sim\mathbb{P}^{[1]}_\beta(\zeta):=\mathrm{PK}_\beta\big(r^{[1]}_{\beta,\zeta}\cdot f_\beta\big)$, where $r^{[1]}_{\beta,\zeta}(t)=\zeta^{1/\beta-1}t\mathrm{e}^{-\zeta^{1/\beta}t}\mathrm{e}^{\zeta}/\beta$.

The case of $m=1$ also allows us to recover the Poisson–Dirichlet coagulation/fragmentation duality results of [31] based on independent $\mathrm{PD}(\alpha,\theta)$ and $\mathrm{PD}(\beta/\alpha,\theta/\alpha)$ distributions for the general case of $\theta>-\beta$. The case of $m=0$ is fairly straightforward but we can recover the duality only for $\theta>0$ using [37, Proposition 21]. Results for general m using Proposition 4.1 are also manageable but require too many additional details for the present exposition.

Let $\big(\tau^{(1)}_\beta(\zeta v),\, v\in[0,1]\big)$ denote the subordinator τ_β run up to a length of ζ, and let $\tau^{(2)}_\beta(G_{(1-\beta)/\beta})\overset{\mathrm{d}}{=}G_{1-\beta}$ denote its total mass over the disjoint interval of length $G_{(1-\beta)/\beta}$, where $G_a\sim\mathrm{Gamma}(a,1)$ for $a>0$ (see [37, Proposition 21]. This implies that $\tau_\beta(\zeta+G_{(1-\beta)/\beta})\overset{\mathrm{d}}{=}\tau^{(1)}_\beta(\zeta)+\tau^{(2)}_\beta(G_{(1-\beta)/\beta})$. Although not well known, it follows from [24] that the corresponding random distribution function (bridge) in the case of $\mathbf{V}\sim\mathbb{P}^{[1]}_\beta(\zeta)$ has the representation, for $v\in[0,1]$,

$$F_{\mathbf{V}}(v)\overset{\mathrm{d}}{=}\frac{\tau_\beta(\zeta v+G_{(1-\beta)/\beta}1_{\{U_1\leq v\}})}{\tau_\beta(\zeta+G_{(1-\beta)/\beta})},\quad \text{with } T_{\mathbf{V}}\overset{\mathrm{d}}{=}\frac{\tau_\beta(\zeta+G_{(1-\beta)/\beta})}{\zeta^{1/\beta}}. \tag{6.1}$$

In this representation, we consider U_1 to be the first atom picked in a sample. This is the concomitant of the mass corresponding to the first size-biased pick, say $\tilde{P}_{1,\beta}(\zeta)$, from $\mathbf{V}$ appearing

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

in the size-biased representation

$$F_{\mathbf{V}}(v) := (1-\tilde{P}_{1,\beta}(\zeta))\frac{\tau_\beta^{(1)}(\zeta v)}{\tau_\beta^{(1)}(\zeta)} + \tilde{P}_{1,\beta}(\zeta)1_{\{U_1\le v\}}, \tag{6.2}$$

where $\tilde{P}_{1,\beta}(\zeta) := \tau_\beta^{(2)}(G_{(1-\beta)/\beta})/\tau_\beta(\zeta + G_{(1-\beta)/\beta})$ and we use $\tau_\beta(\zeta + G_{(1-\beta)/\beta}) := \tau_\beta^{(1)}(\zeta) + \tau_\beta^{(2)}(\gamma_{(1-\beta)/\beta})$. Our purpose in the next result is to determine, given that $\mathbf{V}\sim\mathbb{P}_\beta^{[1]}(\zeta) := \mathrm{PK}_\beta\big(r_{\beta,\zeta}^{[1]}\cdot f_\beta\big)$, explicit equality in distribution constructions of $(F_{\tilde{\mathbf{V}}}, F_{\mathbf{Q}})$ satisfying $F_{\mathbf{V}} = F_{\tilde{\mathbf{V}}}\circ F_{\mathbf{Q}}$, where $\tilde{\mathbf{V}}\overset{\mathrm{d}}{=}\mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V})$ and $\mathbf{Q}$ is the coagulating mass partition. Furthermore, we show how these can be used to recover the duality in the Poisson–Dirichlet case of [31] for all $\theta > -\beta$.

Here, we apply Proposition 4.1, as well as distribution theory connected with these generalized gamma models and the Poisson–Dirichlet family, to identify all the relevant distributions in the next result, which is new.

Proposition 6.1. *Consider the variables* $\mathbf{V}$ *and* $(\tilde{\mathbf{V}},\mathbf{Q})$ *forming the coagulation and fragmentation operations as described in Proposition 4.1, with* $F_{\mathbf{V}} = F_{\tilde{\mathbf{V}}}\circ F_{\mathbf{Q}}$ *and where* $\mathbf{V}\sim\mathbb{P}_\beta^{[1]}(\zeta) := \mathrm{PK}_\beta\big(r_{\beta,\zeta}^{[1]}\cdot f_\beta\big)$*, and thus* $(\tilde{\mathbf{V}},\mathbf{Q})\sim\mathrm{P}_\alpha^{\beta/\alpha}\big(r_{\beta,\zeta}^{[1]}\big)$*, with joint density of* $(T_{\tilde{\mathbf{V}}}, T_{\mathbf{Q}})$ *given by* $r_{\beta,\zeta}^{[1]}(sy^{1/\alpha})f_\alpha(s)f_{\beta/\alpha}(y)$ *with respective arguments (s,y). Then:*

(i) $\mathbf{Q}\sim\mathbb{P}_{\beta/\alpha}^{[1]}(\zeta)$ *where, for* $G_{(\alpha-\beta)/\beta} = G_{(1-\beta/\alpha)/(\beta/\alpha)}\sim\mathrm{Gamma}((1-\beta/\alpha)/(\beta/\alpha), 1)$*,*

$$F_{\mathbf{Q}}(v)\overset{\mathrm{d}}{=}\frac{\tau_{\beta/\alpha}(\zeta v + G_{(\alpha-\beta)/\beta}1_{\{U_1\le v\}})}{\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta})},\quad \text{with } T_{\mathbf{Q}}\overset{\mathrm{d}}{=}\frac{\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta})}{\zeta^{\alpha/\beta}}, \tag{6.3}$$

where $T_{\mathbf{Q}}$ *has density* $r_{\beta/\alpha,\zeta}^{[1]}(y)f_{\beta/\alpha}(y)$*.*

(ii) $\tilde{\mathbf{V}}\mid T_{\mathbf{Q}} = y\sim\mathbb{P}_\alpha^{[1]}(\zeta^{\alpha/\beta}y)$*, hence* $\tilde{\mathbf{V}}\overset{\mathrm{d}}{=}\mathrm{FRAG}_{\alpha,-\beta}(\mathbf{V})\sim\mathbb{P}_\alpha^{[1]}(\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta})) := \int_0^\infty\mathbb{P}_\alpha^{[1]}(\zeta^{\alpha/\beta}y)r_{\beta/\alpha,\zeta}^{[1]}(y)f_{\beta/\alpha}(y)\,\mathrm{d}y$ *and where, for* $G_{(\alpha-\beta)/\beta}$ *independent of* $G_{(1-\alpha)/\alpha}$*,*

$$F_{\tilde{\mathbf{V}}}(u)\overset{\mathrm{d}}{=}\frac{\tau_\alpha\big(\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta})u + G_{(1-\alpha)/\alpha}1_{\{\tilde{U}_1\le u\}}\big)}{\tau_\alpha\big(\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta}) + G_{(1-\alpha)/\alpha}\big)},$$
$$\text{with}\quad T_{\tilde{\mathbf{V}}}\overset{\mathrm{d}}{=}\frac{\tau_\alpha\big(\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta}) + G_{(1-\alpha)/\alpha}\big)}{[\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta})]^{1/\alpha}}. \tag{6.4}$$

(iii) $\mathbf{V}\sim\mathbb{P}_\beta^{[1]}(G_{(\theta+\beta)/\beta}) = \mathrm{PD}(\beta,\theta)$ *for* $\zeta\overset{\mathrm{d}}{=}G_{(\theta+\beta)/\beta}$*,* $\theta > -\beta$*. Hence*

(iv) $\tilde{\mathbf{V}}\sim\mathbb{P}_\alpha^{[1]}(\tau_{\beta/\alpha}(G_{(\theta+\beta)/\beta} + G_{(\alpha-\beta)/\beta})) = \mathrm{PD}(\alpha,\theta)$ *independent of*

$$\mathbf{Q}\sim\mathbb{P}_{\beta/\alpha}^{[1]}(G_{(\theta+\beta)/\beta}) = \mathrm{PD}(\beta/\alpha,\theta/\alpha).$$

Proof. The results follow from an application of Proposition 4.1 using $h(t) = r_{\beta,\zeta}^{[1]}(t)$, the distributional representation of $T_{\mathbf{Q}}$, and the appropriate Gamma randomization to obtain independent PD laws. As mentioned, the generalized gamma subordinator representation of $T_{\mathbf{Q}}$, $T_{\mathbf{V}}$, and Poisson–Dirichlet distributional identities can be found in [21, 24, 25]. For more clarity it is straightforward to show that $r_{\beta/\alpha,\zeta}^{[1]}(y) = \int_0^\infty r_{\beta,\zeta}^{[1]}(sy^{1/\alpha})f_\alpha(s)\,\mathrm{d}s$. As further checks, in verifying (iii) and (iv) we have

$$\mathbb{E}\Big[r^{[1]}_{\beta,G_{(\theta+\beta)/\beta}}(t)\Big] = \frac{\int_0^\infty r^{[1]}_{\beta,x}(t)x^{((\theta+\beta)/\beta)-1}\mathrm{e}^{-x}\,\mathrm{d}x}{\Gamma(\theta/\beta+1)} = \frac{t^{-\theta}}{\mathbb{E}[T_\beta^{-\theta}]}.$$

Set $t = sy^{1/\alpha}$ to obtain the independence. □

Although $F_{\mathbf{V}}(v) = F_{\tilde{\mathbf{V}}}(F_{\mathbf{Q}}(v))$ is a condition of the result, we can directly check that the composition of the subordinator representations in (6.3) and (6.4) results in a $\mathbb{P}^{[1]}_\beta(\zeta) := \mathrm{PK}_\beta\big(r^{[1]}_{\beta,\zeta}\cdot f_\beta\big)$-bridge. A key point is that, by definition, as the concomitant of the first size-biased pick, $U_1 = F^{-1}_{\mathbf{Q}}(\tilde{U}_1)$, where $F^{-1}_{\mathbf{Q}}$ denotes the quantile function of $F_{\mathbf{Q}}$. This results in $\tau_\alpha\big(\tau_{\beta/\alpha}(\zeta + G_{(\alpha-\beta)/\beta})F_{\mathbf{Q}}(v) + G_{(1-\alpha)/\alpha}1_{\{\tilde{U}_1\le F_{\mathbf{Q}}(v)\}}\big)$, written as

$$\tau_\alpha\big(\tau_{\beta/\alpha}\big(\zeta v + G_{(\alpha-\beta)/\beta}1_{\{U_1\le v\}}\big) + G_{(1-\alpha)/\alpha}1_{\{U_1\le v\}}\big) \overset{\mathrm{d}}{=} \tau_\beta\big(\zeta v + G_{(1-\beta)/\beta}1_{\{U_1\le v\}}\big),$$

where the term on the right-hand side involves using the facts discussed around the construction of (6.2) and (6.3), where we can take $\tau_{\beta/\alpha}(G_{(\alpha-\beta)/\beta}) \overset{\mathrm{d}}{=} G_{(\alpha-\beta)/\alpha}$, $\tau_\alpha(G_{(\alpha-\beta)/\alpha} + G_{(1-\alpha)/\alpha}) \overset{\mathrm{d}}{=} G_{1-\beta} \overset{\mathrm{d}}{=} \tau_\beta(G_{(1-\beta)/\beta})$, and $\tau_\alpha(\tau_{\beta/\alpha}(\zeta v)) \overset{\mathrm{d}}{=} \tau_\beta(\zeta v)$. This leads to the representations in (6.1) and (6.2).

Remark 6.2. If $\mathbf{V}_\alpha \sim \mathrm{PD}(\alpha, 0)$ independent of $\mathbf{Q}_{\beta/\alpha} \sim \mathrm{PD}(\beta/\alpha, 0)$, it is evident that

$$(\mathbf{V}_\alpha, \mathbf{Q}_{\beta/\alpha}) \mid N(\zeta^{1/\beta}T_\alpha(T_{\beta/\alpha}(1))) = m \sim \mathrm{P}^{\beta/\alpha}_\alpha\big(r^{[m]}_{\beta,\zeta}\big).$$

Appendix A.

The EPPF of the $\mathrm{PD}_{\alpha|\beta}(\alpha \mid y)$ Gibbs partition of $[n]$ in Theorem 5.1 may be alternatively expressed in terms of Fox-H functions [29] as

$$\frac{\alpha H^{0,2}_{2,2}\Big[y\Big|^{(1-1/\beta,1/\beta),(1-1/\alpha-k,1/\alpha)}_{(1-1/\alpha,1/\alpha),(-n,1)}\Big]}{H^{0,1}_{1,1}\Big[y\Big|^{(1-1/\beta,1/\beta)}_{(0,1)}\Big]}\frac{\Gamma(n)}{\Gamma(k)}p_\alpha(n_1,\ldots,n_k).$$

This expression follows by noting the Fox-H representations for $f_{\beta/\alpha}$ and $f^{(n-k\alpha)}_{\alpha,k\alpha}$, followed by applying [8, Theorem 4.1]. Otherwise the details are similar to the arguments in [20].

Acknowledgement

We thank the anonymous referees for many suggestions to help improve the presentation of this work.

Funding information

L.F. James was supported in part by grants RGC-GRF 16301521, 16300217, and 601712 of the Research Grants Council (RGC) of the Hong Kong SAR.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press

References

[1] BACALLADO, S., BATTISTON, M., FAVARO, S. AND TRIPPA, L. (2017). Sufficientness postulates for Gibbs-type priors and hierarchical generalizations. *Statist. Sci.* **32**, 487–500.

[2] BASDEVANT, A. L. (2006). Ruelleś probability cascades seen as a fragmentation process. *Markov Process. Relat. Fields* **12**, 447–474.

[3] BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes*. Cambridge University Press.

[4] BERTOIN, J. AND LE GALL, J.-F. (2000). The Bolthausen–Sznitman coalescent and the genealogy of continuous-state branching processes. *Prob. Theory Relat. Fields* **117**, 249–266.

[5] BERTOIN, J. AND PITMAN, J. (2000). Two coalescents derived from the ranges of stable subordinators. *Electron. J. Prob.* **5**, 1–17.

[6] BOLTHAUSEN, E. AND SZNITMAN, A. S. (1998). On Ruelleś probability cascades and an abstract cavity method. *Commun. Math. Phys.* **197**, 247–276.

[7] CARON, F. AND FOX, E. B. (2017). Sparse graphs using exchangeable random measures. *J. R. Statist. Soc. B* **79**, 1–44.

[8] CARTER, B. D. AND SPRINGER, M. D. (1977). The distribution of products, quotients and powers of independent H-function variates. *SIAM J. Appl. Math.* **33**, 542–558.

[9] CRANE, H (2016). The ubiquitous Ewens sampling formula. *Statist. Sci.* **31**, 1–19.

[10] CURIEN, N. AND HAAS, B. (2013). The stable trees are nested. *Prob. Theory Relat. Fields* **157**, 847–883.

[11] DE BLASI, P., FAVARO, S., LIJOI, A., MENA, R., PRÜNSTER, I. AND RUGGIERO, M. (2015). Are Gibbs-type priors the most natural generalization of the Dirichlet process? *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 212–229.

[12] DONG, R., GOLDSCHMIDT, C. AND MARTIN, J. (2006). Coagulation–fragmentation duality, Poisson–Dirichlet distributions and random recursive trees. *Ann. Appl. Prob.* **16**, 1733–1750.

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

[14] GASTHAUS, J. AND TEH, Y. W. (2010). Improvements to the sequence memoizer. In *Proc. 24th Int. Conf. Neural Information Processing Systems*, Vol. 1, eds. J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta. ACM, pp. 685–693.

[15] GNEDIN, A. AND PITMAN, J. (2006). Exchangeable Gibbs partitions and Stirling triangles. *J. Math. Sci.* **138**, 5674–5685.

[16] GORENFLO, R., KILBAS, A. A., MAINARDI, F. AND ROGOSIN, S. V. (2014). *Mittag-Leffler Functions, Related Topics and Applications*. Springer, Berlin.

[17] GRIFFITHS, R. C. AND SPANÒ, D. (2007). Record indices and age-ordered frequencies in exchangeable Gibbs partitions. *Electron. J. Prob.* **12**, 1101–1130.

[18] HAAS, B., MIERMONT, G., PITMAN, J. AND WINKEL, M. (2008). Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *Ann. Prob.* **36**, 1790–1837.

[19] HAAS, B., PITMAN, J. AND WINKEL, M. (2009). Spinal partitions and invariance under re-rooting of continuum random trees. *Ann. Prob.* **37**, 1381–1411.

[20] HO, M.-W., JAMES, L. F. AND LAU, J. W. (2007). Gibbs partitions (EPPF's) derived from a stable subordinator are Fox H and Meijer G transforms. Preprint, arXiv:0708.0619 [math.PR].

[21] HO, M.-W., JAMES, L. F. AND LAU, J. W. (2021). Gibbs partitions, Riemann–Liouville fractional operators, Mittag-Leffler functions, and fragmentations derived from stable subordinators. *J. Appl. Prob.* **58**, 314–334.

[22] ISHWARAN, H. AND James. L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* **96**, 161–173.

[23] JAMES, L. F. (2010). Lamperti type laws. *Ann. Appl. Prob.* **20**, 1303–1340.

[24] JAMES, L. F. (2019). Stick-breaking Pitman–Yor processes given the species sampling size. Preprint, arXiv:1908.07186 [math.ST].

[25] JAMES, L. F. (2022). Single-block recursive Poisson–Dirichlet fragmentations of normalized generalized Gamma processes. *Mathematics* **10**, 561.

[26] JAMES, L. F, LEE, J. AND ROSS, N. (2024). Network and interaction models for data with hierarchical granularity via fragmentation and coagulation. Preprint, arXiv:2408.04866 [math.ST].

[27] LEBEDEV, N. N. (1972). *Special Functions and Their Applications*. Dover Publications, New York.

[28] LOMELI, M., FAVARO, S. AND TEH, Y. W. (2017). A marginal sampler for σ-stable Poisson–Kingman mixture models. *J. Comput. Graph. Statist.* **26**, 44–53.

[29] MATHAI, A. M., SAXENA, R. K. AND HAUBOLD, H. J. (2010). *The H-Function. Theory and Applications*. Springer, New York.

[30] PITMAN, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics, Probability and Game Theory* (IMS Lect. Notes Monogr. Ser. **30**). Institute of Mathematical Statistics, Hayward, CA, pp. 245–267.

[31] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Prob.* **27**, 1870–1902.

[32] PITMAN, J. (2003). Poisson–Kingman partitions. In *Statistics and Science: A Festschrift for Terry Speed* (IMS Lect. Notes Monogr. Ser. **40**). Institute of Mathematical Statistics, Beachwood, OH, pp. 1–34.
[33] PITMAN, J. (2006). *Combinatorial Stochastic Processes* (Lect. Notes Math. **1875**). Springer, Berlin.
[34] PITMAN, J. (2017). Mixed Poisson and negative binomial models for clustering and species sampling. Unpublished manuscript.
[35] PITMAN, J. AND YAKUBOVICH, Y. (2018). Ordered and size-biased frequencies in GEM and Gibbs models for species sampling. *Ann. Appl. Prob.* **28**, 1793–1820.
[36] PITMAN, J. AND YOR, M. (1992). Arcsine laws and interval partitions derived from a stable subordinator. *Proc. London Math. Soc*. **65**, 326–356.
[37] PITMAN, J. AND YOR, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Prob.* **25**, 855–900.
[38] WOOD, F., GASTHAUS, J., ARCHAMBEAU, C., JAMES, L. F. AND TEH, Y. W. (2011). The sequence memoizer. *Commun. ACM* **54**, 91–98.

https://doi.org/10.1017/jpr.2025.28 Published online by Cambridge University Press