Statistics Surveys
Vol. 3 (2009) 47–95
ISSN: 1935-7516
DOI: 10.1214/09-SS041

Distributional properties of means of random probability measures*

Antonio Lijoi†

Department of Economics and Quantitative Methods, University of Pavia
via San Felice 5, 27100 Pavia (Italy)
e-mail: lijoi@unipv.it

and

Igor Prünster†

Department of Statistics and Applied Mathematics, University of Turin
Corso Unione Sovietica 218/bis, 10134 Turin (Italy)
e-mail: igor@econ.unito.it

Abstract: The present paper provides a review of the results concerning distributional properties of means of random probability measures. Our interest in this topic has originated from inferential problems in Bayesian Nonparametrics. Nonetheless, it is worth noting that these random quantities play an important role in seemingly unrelated areas of research. In fact, there is a wealth of contributions both in the statistics and in the probability literature that we try to summarize in a unified framework. Particular attention is devoted to means of the Dirichlet process given the relevance of the Dirichlet process in Bayesian Nonparametrics. We then present a number of recent contributions concerning means of more general random probability measures and highlight connections with the moment problem, combinatorics, special functions, excursions of stochastic processes and statistical physics.

AMS 2000 subject classifications: Primary 62F15, 62E15; secondary 60G57.
Keywords and phrases: Bayesian Nonparametrics, completely random measures, Cifarelli–Regazzini identity, Dirichlet process, functionals of random probability measures, generalized Stieltjes transform, neutral to the right processes, normalized random measures, posterior distribution, random means, random probability measure, two–parameter Poisson–Dirichlet process.

Received January 2009.

Contents

*This paper was accepted by Subhashis Ghosal, Associate Editor for the IMS.
†Also affiliated with the Institute of Applied Mathematics and Information Technology, CNR-IMATI, Milan (Italy).
‡Also affiliated with Collegio Carlo Alberto and ICER, Turin, Italy.

1. Introduction

Random probability measures are a fundamental tool in the Bayesian Nonparametrics since their probability distributions identify priors for Bayesian inference. Indeed, let $\mathbb{X}$ be a complete and separable metric space endowed with the Borel σ–algebra $\mathscr{X}$ and denote by $(\mathbb{P}_{\mathbb{X}}, \mathcal{P}_{\mathbb{X}})$ the space of probability distributions defined on $(\mathbb{X}, \mathscr{X})$ equipped with the corresponding Borel σ field. Suppose $(X_n)_{n\geq 1}$ is a sequence of exchangeable random variables defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and taking values in $(\mathbb{X}, \mathscr{X})$. According to de Finetti's representation theorem, there exists a probability measures Π on the space of probability distributions $(\mathbb{P}_{\mathbb{X}}, \mathcal{P}_{\mathbb{X}})$ on $(\mathbb{X}, \mathscr{X})$ such that

$$\mathbb{P}[X_1 \in A_1, \ldots, X_n \in A_n] = \int_{\mathbb{P}_{\mathbb{X}}} \prod_{i=1}^{n} p(A_i)\, \Pi(\mathrm{d}p)$$

for any $A_1, \ldots, A_n$ in $\mathscr{X}$ and $n \geq 1$. Hence, there exists a random probability measure $\tilde{\mathscr{P}}$, defined on $(\mathbb{X}, \mathscr{X})$ and taking values in $(\mathbb{P}_{\mathbb{X}}, \mathcal{P}_{\mathbb{X}})$, conditionally on which the X_i's are independent and identically distributed (i.i.d.) with common distribution $\tilde{\mathscr{P}}$. The distribution of $\tilde{\mathscr{P}}$ coincides with Π and it acts a prior distribution in Bayesian inference. The most celebrated example of prior Π is the Dirichlet process prior introduced by Ferguson in [26].

Within this framework, a natural object of study are the distributional properties of means of random probability measures i.e.

$$\int_{\mathbb{X}} f(x)\tilde{\mathscr{P}}(\mathrm{d}x), \tag{1.1}$$

where f is some real–valued function. This important line of research was initiated in a couple of pioneering papers by Cifarelli and Regazzini. See [8, 9].

Motivated by statistical considerations, they deal with the problem of determining the probability distribution of a generic associative mean which admits the representation

$$M(\tilde{\mathscr{P}};\varphi)=\varphi^{-1}\left(\int_{\mathbb{X}}\varphi(x)\;\tilde{\mathscr{P}}(\mathrm{d}x)\right).$$

where φ is a continuous and increasing real–valued function. The problem, then, reduces to studying the random quantity $\int_{\mathbb{R}} x\,\tilde{\mathscr{P}}(\mathrm{d}x)$ given that $\int_{\mathbb{R}}|x|\,\tilde{\mathscr{P}}(\mathrm{d}x)<\infty$ almost surely. In particular, in [9] the authors introduce a series of tools and techniques that, later in [10], turned out to be fundamental for the determination of the probability distribution of $\int_{\mathbb{X}} f(x)\tilde{\mathscr{P}}(\mathrm{d}x)$ when $\tilde{\mathscr{P}}$ is a Dirichlet process. An appealing aspect of this topic is that distributional results concerning the random quantity in (1.1) are also of interest in research areas not related to Bayesian nonparametric inference. This fact was effectively emphasized in [15] and discussed in several later papers to be mentioned in the following sections.

There has recently been a growing interest in the analysis of distributional properties of random probability measures as witnessed by the variety of publications that have appeared in the statistics and probability literature. Since such an interest originates from different lines of research, we believe it is useful to provide, within an unified framework, an up to date account of the achieved results. The present survey aims at detailing the origin of the various contributions and at pointing out the connections among them. In pursuing this goal, we split the paper into two parts: the first one deals with means of the Dirichlet process, whereas the second part will focus on means of more general random probability measures. In both cases, we will provide, whenever known in the literature, the exact evaluation of the corresponding probability distribution and of other distributional characteristics such as, e.g., the moments. The last section provides some concluding remarks and some possible future lines of research. Finally, the Appendix concisely summarizes a few preliminary notions that play an important role throughout the paper.

2. Means of Dirichlet processes

Before stating the main results, and technical issues, related to the determination of the probability distribution of random Dirichlet means, we briefly recall the notion of a Dirichlet process $\tilde{\mathscr{D}}_\alpha$ with parameter measure α. There are several different constructions of the Dirichlet process: each one has the merit of highlighting a peculiar aspect of this important nonparametric prior. Here below we present four possible definitions of the Dirichlet process, which will be used in the sequel. Similar constructions can also be used to define more general classes of nonparametric priors and the Dirichlet process is typically the only member shared by all these resulting families of random probability measures. Besides the obvious historical reasons, this fact somehow justifies the common view on the Dirichlet process as a cornerstone of Bayesian Nonparametrics.

Following Ferguson's original definition in [26], one can construct the Dirichlet process $\tilde{\mathscr{D}}_\alpha$ in terms of a consistent family of finite–dimensional Dirichlet

distributions. Recall that the Dirichlet distribution, $D_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$, is the probability distribution on the $(k-1)$–dimensional simplex $\Delta_{k-1} := \{(x_1, \ldots, x_{k-1}) : x_1 \geq 0, \ldots, x_{k-1} \geq 0, \sum_{i=1}^{k-1} x_i \leq 1\}$ given by

$$D_{\boldsymbol{\alpha}}(\mathrm{d}x_1, \ldots, \mathrm{d}x_{k-1}) = \frac{\Gamma(\theta)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \left(\prod_{i=1}^{k-1} x_i^{\alpha_i - 1} \right) (1 - |\boldsymbol{x}|)^{\alpha_k} \, \mathrm{d}x_1 \ldots \mathrm{d}x_{k-1}$$

where $|\boldsymbol{x}| := x_1 + \cdots + x_{k-1}$. Let α be a finite measure on $\mathbb{X}$. A random probability measure is termed *Dirichlet process*, $\tilde{\mathscr{D}}_\alpha$ if for any $k \geq 1$ and any measurable partition $A_1, \ldots, A_k$ of $\mathbb{X}$ one has

$$(\tilde{\mathscr{D}}_\alpha(A_1), \ldots, \tilde{\mathscr{D}}_\alpha(A_k)) \sim D_{\boldsymbol{\alpha}}$$

with $\boldsymbol{\alpha} = (\alpha(A_1), \ldots, \alpha(A_k))$. The parameter measure can obviously be decomposed as $\alpha(\,\cdot\,) = \theta \, P_0(\,\cdot\,)$, where $\theta > 0$ is the total mass of α and $P_0 = \alpha/\theta$ is a probability distribution on $(\mathbb{X}, \mathscr{X})$.

An alternative definition of the Dirichlet process, pointed out in [26], makes use of completely random measures (CRMs). See the Appendix for a concise account and some noteworthy examples of CRMs. Indeed, if $\tilde{\gamma}$ is a gamma CRM on $\mathbb{X}$ with finite parameter measure α, then it can be shown that the Dirichlet process coincides, in distribution, with the normalized gamma process

$$\frac{\tilde{\gamma}(\,\cdot\,)}{\tilde{\gamma}(\mathbb{X})} \stackrel{d}{=} \tilde{\mathscr{D}}_\alpha(\,\cdot\,). \tag{2.1}$$

As we shall see this definition suggests interesting generalizations of the Dirichlet process and is useful for investigating the probability distribution of the corresponding random means. See Section 3.2.

Another construction which has recently become very popular in Bayesian Nonparametrics is based on a stick–breaking procedure. This construction was originally proposed in [78] and it can be described as follows. Let $(V_i)_{\geq 1}$ be a sequence of independent and identically distributed (i.i.d.) random variables, with $V_i \sim \mathrm{beta}(1, \theta)$ and define random probability weights $(\tilde{p}_j)_{j \geq 1}$ as

$$\tilde{p}_1 = V_1, \qquad \tilde{p}_j = V_j \prod_{i=1}^{j-1} (1 - V_i) \quad j = 2, 3, \ldots \,. \tag{2.2}$$

If $(Y_i)_{i \geq 1}$ is a sequence of i.i.d. random variables whose common probability distribution is P_0, then

$$\sum_{j \geq 1} \tilde{p}_j \, \delta_{Y_j}(\,\cdot\,) \stackrel{d}{=} \tilde{\mathscr{D}}_\alpha(\,\cdot\,). \tag{2.3}$$

with $\alpha = \theta P_0$. This construction can be generalized to define interesting classes of random probability measures among which it is worth mentioning the two–parameter Poisson–Dirichlet process as a remarkable example. See Section 3.1.

Finally, we mention a definition which stems from applications to survival analysis. In this setting one has $\mathbb{X} = \mathbb{R}^+$ since the object of investigation are

survival times. Suppose, for simplicity, that P_0 is a non–atomic measure and let $S_0(x) = P_0((x,\infty))$ for any x. If $\tilde{\mu}$ is a CRM with Lévy intensity

$$\nu(\mathrm{d}s, \mathrm{d}x) = \mathrm{e}^{-s\,\theta\,S_0(x)} \left\{1 - \mathrm{e}^{-s}\right\}^{-1} \alpha(\mathrm{d}x)\,\mathrm{d}s,$$

and $\tilde{\mathscr{P}}$ is a random probability measure on $\mathbb{R}^+$ characterized by the cumulative distribution function $\tilde{F}(t) = 1 - \mathrm{e}^{-\tilde{\mu}((0,t])}$, then $\tilde{\mathscr{P}}(\,\cdot\,) \stackrel{d}{=} \tilde{\mathscr{D}}_\alpha(\,\cdot\,)$. This provides a representation of the Dirichlet process as a neutral to the right process. See [17] and [27].

Each of the definitions we have briefly described are useful for determining distributional results about a Dirichlet random mean $\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x)$, where f is a real–valued measurable function defined on $\mathbb{X}$. Indeed, the most convenient construction of the underlying Dirichlet process to resort to is suggested by the specific technique one applies for investigating the properties of $\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x)$. Before entering the main details, it should be recalled that an important preliminary step requires the specification of conditions ensuring finiteness of $\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x)$. In other words, one has to determine those measurable functions f for which the integral $\int_{\mathbb{X}} |f(x)|\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x)$ is almost surely finite. As shown in [25] a necessary and sufficient condition is given by

$$\int_{\mathbb{X}} \log(1 + |f(x)|)\,\alpha(\mathrm{d}x) < \infty\,. \tag{2.4}$$

An alternative proof of this fact can be found in [12]. Let now $\alpha_f = \alpha \circ f^{-1}$ denote the image measure of α through f and note that, since

$$\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x) \stackrel{d}{=} \int_{\mathbb{R}} x\,\tilde{\mathscr{D}}_{\alpha_f}(\mathrm{d}x),$$

with no loss of generality we can confine ourselves to considering distributional properties related to $M(\tilde{\mathscr{D}}_\alpha) := \int_{\mathbb{R}} x\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x)$. Moreover, denote by $\mathscr{M}$ the set of non null, non degenerate and finite measures on $\mathbb{R}$ and define the following classes of measures indexed by the total mass parameter $\theta \in (0,\infty)$

$$\mathbb{A}_\theta := \left\{\alpha \in \mathscr{M} :\ \alpha(\mathbb{R}) = \theta\ \wedge\ \int_{\mathbb{R}} \log(1 + |x|)\,\alpha(\mathrm{d}x) < \infty\right\} \tag{2.5}$$

Throughout the present Section will provide a broad description of the main results achieved in the literature about the representation of the probability distribution $Q(\,\cdot\,;\alpha)$ defined by

$$Q(B;\alpha) = \mathbb{P}\left[\int_{\mathbb{X}} x\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x) \in B\right] \qquad \forall B \in \mathscr{B}(\mathbb{R}). \tag{2.6}$$

Note that, if not otherwise stated, it will be tacitely assumed that the Dirichlet base measure α is an element of $\mathbb{A}_\theta$ for some $\theta > 0$.

2.1. Cifarelli–Regazzini identity and related results

A first remarkable result about random Dirichlet means is an identity which was obtained by Cifarelli and Regazzini [9, 10]. This provides a representation of the generalized Stieltjes transform of order θ

$$\mathcal{S}_\theta[z; M(\tilde{\mathscr{D}}_\alpha)] := \mathbb{E}\left[\left(z + M(\tilde{\mathscr{D}}_\alpha)\right)^{-\theta}\right] \tag{2.7}$$

of the law of the mean $M(\tilde{\mathscr{D}}_\alpha)$ in terms of the Laplace functional of a gamma process. It should be recalled that the order $\theta > 0$ coincides with the total mass of the parameter measure α, i.e. $\theta = \alpha(\mathbb{R})$. The interest in (2.7) is motivated by the fact that one can resort to inversion formulae of generalized Stieltjes transforms and deduce a representation of the probability distribution of $M(\tilde{\mathscr{D}}_\alpha)$. See Appendix B.1 for a concise account on generalized Stieltjes transforms. In the following $\mathrm{Re}\, z$ and $\mathrm{Im}\, z$ stand for the real and imaginary parts of $z \in \mathbb{C}$, respectively. The important identity is then as follows.

Theorem 2.1. (Cifarelli-Regazzini identity) *For any* $z \in \mathbb{C}$ *such that* $\mathrm{Im}(z) \neq 0$, *one has*

$$\mathcal{S}_\theta[z; M(\tilde{\mathscr{D}}_\alpha)] = z^{-\theta} \mathrm{e}^{-\int_{\mathbb{R}} \log[1+z^{-1}x]\, \alpha(\mathrm{d}x)}. \tag{2.8}$$

It is worth noting that the identity (2.8) holds true when the order of the generalized Stieltjes transform coincides with the total mass θ of the parameter measure of the underlying Dirichlet process. An interesting case of (2.8) is obtained when $\theta = 1$ since it reduces to

$$\int_{\mathbb{R}} \frac{1}{z+x}\, Q(\mathrm{d}x; \alpha) = \frac{1}{z}\, \mathrm{e}^{-\int_{\mathbb{R}} \log[1+z^{-1}x]\, P_0(\mathrm{d}x)}. \tag{2.9}$$

Indeed, in the terminology of [49], one says that the probability distribution $Q(\,\cdot\,; \alpha)$ of $M(\tilde{\mathscr{D}}_\alpha)$ is the *Markov transform* of P_0. And the moment problem for $Q(\,\cdot\,; \alpha)$ is determined if and only if it is determined for P_0. See Corollary 3.2.5 in [49]. As one can see from [49] and [15], the interest in (2.8) and (2.9) also arises in other seemingly unrelated areas of research such as, for example, the growth of Young diagrams or the exponential representations of functions of negative imaginary parts.

The proof of (2.8) can be achieved by means of different techniques. In [10] the result is obtained by resorting to analytic arguments whose preliminary application can also be found in [9]. Indeed, the authors consider the truncated functional $U(\tau, T) := \int_{(\tau,T]} x\, \tilde{\mathscr{D}}_\alpha(\mathrm{d}x) - \tau$ with $-\infty < \tau < T < \infty$. Letting $Q_{\tau,T}(\,\cdot\,; \alpha)$ denote the probability distribution of $U(\tau, T)$, they obtain the following series expansion

$$s^{-\theta}\mathcal{S}_\theta(1/s; \tau, T) = \int_{\mathbb{R}^+} \frac{1}{(1+sx)^\theta}\, Q_{\tau,T}(\mathrm{d}x; \alpha) = \sum_{n\geq 0} c_n(\tau, T) s^n \tag{2.10}$$

for any $s \in (0, (T-\tau)^{-1}]$. Differentiating the series in (2.10), term by term, with respect to τ it can be shown that the function $s^{-\theta}\mathcal{S}_\theta(1/s; \tau, T)$ satisfies a partial

differential equation whose solution is $\exp\{-\int_\tau^T [z\,\alpha((x,\infty))](s+x-\tau)^{-1}\,\mathrm{d}x\}$. This concise description of the main ideas hides the technical difficulties one has to overcome in order to get to the final result. The interested reader should check [9] and [10] to appreciate their complete line of reasoning.

An approach based on combinatorial arguments has been pursued in [84] and in [15], under the additional assumption that $\mathrm{supp}(\alpha) = [0,1]$. The result in [84] is more general since it allows to obtain similar identities for the two–parameter Poisson–Dirichlet process to be examined in Section 3.1. Indeed, in [87] one finds a relation between the moments of $Q(\,\cdot\,;\alpha)$ and those of α

$$\int_{[0,1]} x^n\, Q(\mathrm{d}x;\alpha) = \frac{1}{(\theta)_n} \sum_{k=1}^n \theta^k \times \\ \times \sum_{(m_1,\ldots,m_n)\in Z(n,k)} \frac{n!}{\prod_{j\geq 1} m_j! j^{m_j}} \prod_{j=1}^k r_{j,\alpha}^{m_j} \tag{2.11}$$

where $r_{j,\alpha} := \int_{[0,1]} x^j\,\alpha(\mathrm{d}x)$, $(a)_n = \prod_{j=1}^n (a+j-1)$ is the Pochhammer symbol with the convention that $(a)_0 \equiv 1$ and

$$Z(n,k) := \{(m_1,\ldots,m_n)\in\mathbb{Z}_+^n : \sum_i i\,m_i = n,\ \sum_i m_i = k\} \tag{2.12}$$

with $\mathbb{Z}_+ := \{0,1,2,\ldots\}$. Note that $Z(n,k)$ identifies all possible partitions of the set of integers $\{1,\ldots,n\}$ into k sets. It is apparent from its definition that, if $(m_1,\ldots,m_n)\in Z(n,k)$, then $\max\{i : m_i \neq 0\} \leq n-k+1$. The relation in (2.11) can be found, with different proofs, also in other papers as we shall see when discussing moments of random Dirichlet means. Hence, one has

$$H(z) := \int_{\mathbb{R}} \frac{1}{(1-zx)^\theta}\, Q(\mathrm{d}x;\alpha) - 1 = \sum_{n\geq 1} \frac{z^n}{n!}\,(\theta)_n \int_{[0,1]} x^n\, Q(\mathrm{d}x;\alpha) \\ = \sum_{n\geq 1} \frac{z^n}{n!} \sum_{k=1}^n \theta^k \sum_{(m_1,\ldots,m_n)\in Z(n,k)} \frac{n!}{\prod_{j\geq 1} m_j! j^{m_j}} \prod_{j\geq 1} r_{j,\alpha}^{m_j} = \sum_{n\geq 1} \frac{z^n}{n!}\, H_n$$

where H_n is expressed in terms of Bell polynomials as follows

$$H_n = \sum_{k=1}^n f_k B_{n,k}(g) = \sum_{k=1}^n \theta^k \sum_{(m_1,\ldots,m_n)\in Z(n,k)} \frac{n!}{\prod_{j\geq 1} m_j! (j!)^{m_j}} \prod_{j\geq 1} \{g_j\}^{m_j}\,.$$

In the above expression, one has $f_k = \theta^k$ and $g_j = (j-1)! r_{j,\alpha}$ for any $k \geq 1$ and $j \geq 1$. If we define $f(t) = \sum_{n\geq 1} f_n t^n/n! = \exp\{\theta t\} - 1$ and $g(t) = \sum_{n\geq 1} g_n t^n/n! = -\int \log(1-tx)\,\alpha(\mathrm{d}x)$, a result from the theory of symmetric functions yields

$$H(z) = f(g(z)) = \mathrm{e}^{-\theta \log[1-zx]\,\alpha(\mathrm{d}x)} - 1\,.$$

A similar argument is used in [15] where a series expansion of $z^{-\theta}\,\mathcal{S}_\theta[z^{-1}; M(\tilde{\mathscr{D}}_\alpha)]$ along with the so-called *exponential formula* (see, e.g., Corollary 5.1.9 in [81])

$$\sum_{j\geq 0} z^{-j}\mathbb{E}\left[M^j(\tilde{\mathscr{D}}_\alpha)\right] = \mathrm{e}^{-\sum_{j\geq 1} z^{-j} r_{j,\alpha}/j}$$

yield (2.8).

In [28] a simple procedure for deriving $\mathcal{S}_\theta[z; M(\tilde{\mathscr{D}}_\alpha)]$ is proposed. In particular, starting from the case of α with bounded support, a first–type Volterra equation for the Laplace transform m of $M(\tilde{\mathscr{D}}_\alpha)$ is introduced: it then follows that the Laplace transform of $x^{\theta-1}m(x)$, which corresponds to $\mathcal{S}_\theta[z; M(\tilde{\mathscr{D}}_\alpha)]$, satisfies a certain first order ordinary differential equation, whose solution is explicitly given and coincides with (2.8). Then, the result is also extended to the case of α having support bounded from below.

An even simpler proof has been devised in [85] and it relies on the definition of the Dirichlet process in (2.1) as a normalized gamma process. Indeed, using the independence between $\tilde{\gamma}(\mathbb{R})$ and $\tilde{\mathscr{D}}_\alpha$ one has

$$\mathbb{E}\Big[\mathrm{e}^{-z\int_{\mathbb{R}} x\,\tilde{\gamma}(\mathrm{d}x)}\Big] = \mathbb{E}\Big[\mathrm{e}^{-z\,\tilde{\gamma}(\mathbb{R})\,M(\tilde{\mathscr{D}}_\alpha(\mathrm{d}x))}\Big] = \mathbb{E}\Big[\big(1 - zM(\tilde{\mathscr{D}}_\alpha)\big)^{-\theta}\Big]$$

where the last equality follows from the fact that $\tilde{\gamma}(\mathbb{R})$ has the gamma density $g(x) = [\Gamma(\theta)]^{-1}x^{\theta-1}\,\mathrm{e}^{-x}\mathbb{1}_{(0,\infty)}(x)$. See Theorem 2 in [85].

We finally mention that (2.8) can be shown by means of a discretization procedure which consists in evaluating the Stieltjes transform of a Dirichlet random mean $\int x\,\tilde{\mathscr{D}}_{\alpha_m}(\mathrm{d}x)$ with $\alpha_m = \sum_{i=1}^{k_m}\alpha(A_{m,i})\,\delta_{x_{m,i}}$ being a suitable discretization of α. Moreover, α_m converges weakly to α as $m\to\infty$. This approach has been developed in [71] and later applied in [72]. A similar discretization procedure is also at the basis of the proof given in [36].

It should be remarked that in (2.8) the order of the Stieltjes transform on the right–hand side must coincide with the total mass of the parameter measure α. If interest relies in evaluating the probability distribution of $M(\tilde{\mathscr{D}}_\alpha)$, it would be desirable, regardless of the value of θ, to work with the Stieltjes transform of order 1. This is due to the fact that inversion formulae for Stieltjes transforms of order different from 1 are much harder to implement. Extensions of (2.8) where the order of the generalized Stieltjes transform can differ from the total mass $\theta = \alpha(\mathbb{R})$ are given in [59] and [41]. In [59], the authors make use of the connection between Lauricella multiple hypergeometric functions and the Dirichlet distribution. Before recalling this connection, set $\Delta_n = \{(x_1,\ldots,x_n) : x_1\geq 0,\ldots,x_n\geq 0,\ \sum_{i=1}^n x_i\leq 1\}$ and let $\boldsymbol{\alpha} = (b_1,\ldots,b_{n+1})\in(0,\infty)^{n+1}$, with $\theta = \sum_{i=1}^{n+1} b_i$. One then has the power series representation

$$\begin{aligned} F_{_D}(\theta, b_1,\ldots,b_n; c; x_1,\ldots,x_n) &= \sum_{l_1,\ldots,l_n} \frac{(\theta)_{l_1+\cdots+l_n}}{(c)_{l_1+\cdots+l_n}} \prod_{i=1}^n \frac{(b_i)_{l_i}}{l_i!}\, x_i^{l_i} \\ &= \int_{\Delta_n} \frac{D_{\boldsymbol{\alpha}}(\mathrm{d}x_1,\ldots,\mathrm{d}x_n)}{\{1-\sum_{i=1}^n p_i x_i\}^\theta} \end{aligned} \tag{2.13}$$

where F_D denotes the fourth Lauricella hypergeometric function. See (2.3.5) in [23]. Another useful integral representation is

$$F_D(\theta, b_1, \ldots, b_n; c; x_1, \ldots, x_n) = \frac{\Gamma(\theta)}{\Gamma(c)\,\Gamma(\theta - c)} \int_0^1 u^{c-1}(1-u)^{\theta-c-1} \prod_{i=1}^n (1 - ux_i)^{-b_i}\, \mathrm{d}u \quad (2.14)$$

whenever $\theta > 0$ and $\theta - c > 0$. In the case where $\theta = c$ the identity in (2.14) reduces to

$$F_D(\theta, b_1, \ldots, b_n; \theta; x_1, \ldots, x_n) = \prod_{i=1}^n (1 - x_i)^{-b_i}. \quad (2.15)$$

Letting $\alpha = \sum_{i=1}^{n+1} b_i \delta_{x_i}$ be a measure with point masses at $x_1, \ldots, x_n$ and $x_{n+1} = 0$, from (2.13) and (2.15) one finds that

$$(-1)^\theta\, \mathcal{S}_\theta[-1;\, M(\tilde{\mathscr{D}}_\alpha)] = \mathrm{e}^{-\int_{\mathbb{R}} \log[1-x]\, \alpha(\mathrm{d}x)}.$$

Hence, one can easily recover a simplified form of the identity (2.8) from integral representations of F_D. Let us now consider a measure α such that $\alpha(\mathbb{R}) = \theta \in (0, \infty)$ and with $\mathrm{supp}(\alpha)$ contained in a bounded interval $[0, t)$, for some $t > 0$. Using the above mentioned connection between the law of $M(\tilde{\mathscr{D}}_\alpha)$ and F_D, in [71] it has been proved that for any $\theta \geq c$ and $z \in \mathbb{C}$ such that $\mathrm{Im}(z) \neq 0$ when $\mathrm{Re}(z) \leq 0$

$$\mathcal{S}_c[z; M(\tilde{\mathscr{D}}_\alpha)] = z^{\theta-c} \int_0^1 \mathrm{e}^{-\int_{\mathbb{R}} \log[z+ux]\, \alpha(\mathrm{d}x)}\, B(\mathrm{d}u; c, \theta - c) \quad (2.16)$$

where $B(\mathrm{d}u; a, b)$ is the beta probability distribution if $b > 0$ and it reduces to the point mass δ_1 at 1 if $b = 0$. Hence, one obtains a representation of the order c Stieltjes transform of $M(\tilde{\mathscr{D}}_\alpha)$, the order c being not greater than the total mass θ of α, under the assumption that the support α is included in a bounded interval of $\mathbb{R}^+$. Accordingly, if $\theta > 1$, one can fix $c = 1$ and invert (2.16) to obtain a representation for the density function of the mean $M(\tilde{\mathscr{D}}_\alpha)$. Finally, in [59] the authors obtain an extension to the case α being an arbitrary member of $\mathbb{A}_\theta$ as defined in (2.5) and c being any positive number. In particular, using arguments similar to those exploited in [71] it has been shown that

$$\mathcal{S}_c[z; M(\tilde{\mathscr{D}}_\alpha)] = z^{-\theta} \int_0^1 \mathrm{e}^{-\int_{\mathbb{R}} \log[1+z^{-1}ux]\, \alpha(\mathrm{d}x)}\, B(\mathrm{d}u; c, \theta - c) \quad (2.17)$$

if $\theta \in [c, \infty)$, whereas

$$\mathcal{S}_c[z; M(\tilde{\mathscr{D}}_\alpha)] = z^{-\theta}\, [1 - \mathrm{e}^{2\pi(\theta-c)}]^{-1}\, \frac{\Gamma(c)}{\Gamma(\theta)\Gamma(c-\theta)} \times \\ \times \int_0^{(1+)} \mathrm{e}^{-\int_{\mathbb{R}} \log[1+z^{-1}wx]\, \alpha(\mathrm{d}x)}\, w^{\theta-1}(1-w)^{c-\theta-1}\, \mathrm{d}w \quad (2.18)$$

if $\theta \in (0, c)$, and the above integral is evaluated along the path, in the complex plane, starting at $w = 0$, encircling in a counterclockwise sense $w = 1$ and ending at $w = 0$. Hence, the new identities (2.17) and (2.18) express the Stieltjes transform of order c as a mixture of Laplace functional transforms of the gamma process, the mixing measure being a beta probability distribution on $(0, 1)$ or on the complex plane according as to whether c is greater than or less than the total mass θ of the parameter measure α. Given that they have been directly deduced from integral representations of F_D, the authors in [59] have termed them *Lauricella identities*. It should be recalled that this connection between random Dirichlet means and multiple hypergeometric functions has also been exploited in [50] to deduce a multivariate version of (2.8) involving a vector of means $(\int f_1 \, \mathrm{d}\tilde{\mathscr{D}}_\alpha, \ldots, \int f_d \, \mathrm{d}\tilde{\mathscr{D}}_\alpha)$, with $f_1, \ldots, f_d$ being measurable functions such that $\int \log[1 + |f_i|] \, \mathrm{d}\alpha < \infty$ for any $i = 1, \ldots, d$.

Finally, the extension provided in [41] is achieved by introducing a new process taking values on the space of measures on $(\mathbb{X}, \mathscr{X})$. To this end, let $\mathbb{P}_{\tilde{\gamma}}$ denote the probability distribution of a gamma process $\tilde{\gamma}$ with parameter measure α such that $\alpha(\mathbb{X}) = \theta$. Next define, for any $d \in (0, \theta)$, another probability measure $\mathbb{P}_{\tilde{\gamma},d}$ on the space of measures on $(\mathbb{X}, \mathscr{X})$ such that $\mathbb{P}_{\tilde{\gamma},d} \ll \mathbb{P}_{\tilde{\gamma}}$ and

$$\frac{\mathrm{d}\mathbb{P}_{\tilde{\gamma},d}}{\mathrm{d}\mathbb{P}_{\tilde{\gamma}}}(\mu) = \frac{\Gamma(\theta)}{\Gamma(\theta - d)} \, (\mu(\mathbb{X}))^{-d}.$$

The random measure $\tilde{\mu}$ whose distribution coincides with $\mathbb{P}_{\tilde{\gamma},d}$ is named a *Beta-gamma process* by [41]. For any $c > 0$ it is shown that

$$\mathcal{S}_c[z; M(\tilde{\mathscr{D}}_\alpha)] = \mathbb{E}_{\tilde{\gamma},\theta-c}\left[\mathrm{e}^{-z\int_{\mathbb{R}} x \, \mathrm{d}\tilde{\mu}}\right]. \tag{2.19}$$

When $\theta - c > 0$, (2.19) reduces to the Lauricella identity (2.17). On the other hand, when $\theta - c < 0$, (2.19) provides an interesting alternative expression for $\mathcal{S}_\theta$ which can be compared to (2.18).

2.2. Probability distribution of random Dirichlet means

An expression for the Stieltjes transform of $M(\tilde{\mathscr{D}}_\alpha)$, as discussed in the previous section, represents a very useful tool in order to evaluate the probability distribution of $M(\tilde{\mathscr{D}}_\alpha)$. Indeed, one can resort to an inversion formula for $\mathcal{S}_\theta$ and recover $Q(\,\cdot\,; \alpha)$. Before dealing with the subject, it must be recalled that $Q(\,\cdot\,; \alpha)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and we will denote by $q(\,\cdot\,; \alpha)$ the corresponding density function. See Lemma 2 in [10]. Moreover, denote by A the distribution function corresponding to the base measure α, i.e.

$$A(x) = \alpha((-\infty, x]) \qquad \text{for any } x \in \mathbb{R}, \tag{2.20}$$

where obviously $\lim_{x\to\infty} A(x) = \theta$. As outlined in the Appendix, if appropriate conditions are satisfied one can invert $\mathcal{S}_\theta$ for any $\theta > 0$. Nonetheless, as already mentioned, one can more easily handle the inversion formula when $\alpha(\mathbb{R}) = \theta = 1$.

Theorem 2.2. *Let $\alpha \in \mathbb{A}_\theta$ with $\theta = 1$. Then*

$$q(y;\alpha) = \frac{1}{\pi}\, \mathrm{e}^{-\int_{\mathbb{R}} \log|x-y|\, \alpha(\mathrm{d}x)}\, \sin\left(\pi A(y)\right). \tag{2.21}$$

for any y point of continuity of A.

The representation in (2.21) follows from Theorem 1(ii) in [10] under the assumption that, given $\inf \operatorname{supp}(\alpha) =: \tau \in \mathbb{R}$, one has $A(\tau) = 0$. The same representation has been achieved in Proposition 9(iii) of [72] by removing this condition on α. The proof is based on a different approach and makes use of a simple and useful idea given in [33]. Indeed, from (2.1) one has that

$$\mathbb{F}(y) := \mathbb{P}\left[\int_{\mathbb{R}} x\, \tilde{\mathscr{D}}_\alpha(\mathrm{d}x) \le y\right] = \mathbb{P}\left[\int_{\mathbb{R}} (x-y)\, \tilde{\gamma}(\mathrm{d}x) \le 0\right] \tag{2.22}$$

for any $y \in \mathbb{R}$, where we recall that $\tilde{\gamma}$ stands for the gamma process. Note that such an equality will be extensively used later in Section 3.2 for studying means of more general random probability measures. Given this, an application of the inversion formula for characteristic functions in [31] yields an expression for the cumulative distribution function in (2.22) and, consequently, for the corresponding density function $q(\,\cdot\,;\alpha)$.

Unfortunately, the evaluation of $q(\,\cdot\,;\alpha)$ becomes more difficult if $\theta \neq 1$. In [10] one can find an expression for $q(\,\cdot\,;\alpha)$ of $M(\tilde{\mathscr{D}}_\alpha)$ when $\alpha(\mathbb{R}) = \theta > 1$ and the possible jumps of A, defined as in (2.20), are smaller than one and with the additional assumption that the support of α is bounded from below by some $\tau \in \mathbb{R}$ and $A(\tau) = 0$. This assumption is dropped by [72] and [59] to prove the following

Theorem 2.3. *Suppose $\alpha \in \mathbb{A}_\theta$ with $\theta > 1$ and that A defined in (2.20) has jumps of size smaller than one. Then*

$$q(y;\alpha) = \frac{\theta-1}{\pi} \int_{-\infty}^{y} (y-x)^{\theta-2} \mathrm{e}^{-\int_{\mathbb{R}} \log|x-s|\, \alpha(\mathrm{d}s)}\, \sin(\pi\, A(x))\, \mathrm{d}x \tag{2.23}$$

for any point y of continuity for A.

Explicit expressions for $q(\,\cdot\,;\alpha)$ for values of $\theta > 1$ with A having discontinuities with jumps of size greater than or equal to one and for values of $\theta \in (0,1)$ can be found in items (i) and (iv) of Proposition 9 of [72] and in Section 6 of [59].

As far as the evaluation of the cumulative distribution function $\mathbb{F}(\,\cdot\,)$ of $M(\tilde{\mathscr{D}}_\alpha)$ is concerned, from Proposition 3 in [72] one has the following result

Theorem 2.4. *Let $\alpha \in \mathbb{A}_\theta$. Then, for any y in $\mathbb{R}$,*

$$\mathbb{F}(y) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{1}{t}\, \mathrm{e}^{-\frac{1}{2}\int_{\mathbb{R}} \log[1+t^2(y-x)^2]\, \alpha(\mathrm{d}x)} \times \\ \times\, \sin\left(\int_{\mathbb{R}} \arctan\left[t(y-x)\right]\, \alpha(\mathrm{d}x)\right) \mathrm{d}t\,. \tag{2.24}$$

Here below we describe a few examples mainly based on an application of Theorem 2.2 and Theorem 2.3.

Example 2.1. Suppose $\alpha(\mathrm{d}x) = \mathbb{1}_{(0,1)}(x)\,\mathrm{d}x$. Then, from (2.21) it can be easily seen that

$$q(y;\alpha) = \frac{\sin(\pi y)}{\pi}\; y^{-y}\;(1-y)^{-1+y}\;\mathbb{1}_{(0,1)}(y).$$

See also [15]. □

Example 2.2. Let $\alpha(\mathrm{d}x) = \lambda\,\mathrm{e}^{-\lambda x}\mathbb{1}_{(0,\infty)}(x)\,\mathrm{d}x$. Since $\alpha(\mathbb{R}) = 1$, one can apply (2.21) and obtain

$$q(y;\alpha) = \frac{\sin\left(\pi - \pi \mathrm{e}^{-\lambda y}\right)}{\pi\, y}\,\mathrm{e}^{h(y,\lambda)}\;\mathbb{1}_{(0,\infty)}(y)$$

where $h(y,\lambda) = \mathrm{e}^{-\lambda y}\,\mathrm{Ei}(\lambda y)$. Recall Ei is the usual notation for the exponential integral, i.e. $\mathrm{Ei}(x) = \int_{-\infty}^{x} t^{-1}\,\mathrm{e}^{t}\,\mathrm{d}t$ for $x < 0$ and $\mathrm{Ei}(x) = \mathrm{PV}\int_{-\infty}^{x} t^{-1}\,\mathrm{e}^{t}\,\mathrm{d}t$ for $x > 0$, where PV denotes the principal value integral. See also [72]. □

Example 2.3. Suppose $\alpha = \theta_0\delta_0 + \theta_1\delta_1$, with $\theta_0 + \theta_1 = \theta$. If one observes that in this case $M(\tilde{\mathscr{D}}_\alpha) \stackrel{d}{=} \tilde{\mathscr{D}}_\alpha(\{1\})$, then it can be deduced that $q(\cdot,;\alpha)$ is the beta density with parameters (θ_0,θ_1). On the other hand, this same result can also be achieved by resorting to (2.21) and to (2.23) when $\theta = 1$ or $\theta > 1$, respectively. The case $\theta < 1$ can be dealt with by means of the expression provided in Proposition 9(iv) of [72]. □

Example 2.4. Let $\alpha(\mathrm{d}x) = \sigma\pi^{-1}[1+\sigma^2(x-\mu)^2]^{-1}\,\mathrm{d}x$ be the Cauchy density with parameters (μ,σ). It can then be seen that

$$\int_{\mathbb{R}} \log|x-y|\;\alpha(\mathrm{d}x) = \frac{1}{2}\,\log\frac{1+\sigma^2(y-\mu)^2}{\sigma^2}$$

for any $y \in \mathbb{R}$. From (2.21) one has

$$q(y,\alpha) = \frac{\sigma}{\pi\sqrt{1+\sigma^2(y-\mu)^2}}\;\sin\left(\frac{\pi}{2} + \arctan(\sigma(y-\mu))\right) = \frac{\alpha(\mathrm{d}y)}{\mathrm{d}y}.$$

Hence, in this case $Q(\,\cdot\,;\alpha) = \alpha(\,\cdot\,)$ and such a result was first proved in [89]. Furthermore, it can be shown that $Q(\,\cdot\,;\alpha) = \alpha(\,\cdot\,)$ if and only if α is Cauchy. See [10] and [59]. □

2.3. Determination of the parameter measure α

An interesting question to be raised concerns the connection between $q(\,\cdot\,;\alpha)$ and the parameter α. Indeed, one might wonder whether there is a one–to–one correspondence between the set of possible probability distributions of random Dirichlet means and the space of parameter measures α given by $\mathbb{A}_\theta$ in (2.5). Hence, denote by $\mathbb{M}_\theta$ the family of all possible density functions of $M(\tilde{\mathscr{D}}_\alpha)$ as α varies in $\mathbb{A}_\theta$. Indeed, the answer is positive and the identity (2.8) sets up a bijection between the set $\mathbb{A}_\theta$ and $\mathbb{M}_\theta$ as shown, e.g. in [59] and in [36].

Given this important result, one might try to provide an answer to the following question: is it possible to determine the parameter measure $\alpha \in \mathbb{A}_\theta$ yielding a specific distribution in $\mathbb{M}_\theta$? This corresponds to reversing the issue we have been addressing up to now. An early answer can be found in [11]. The authors provided a general solution to the problem, valid for parameter measures whose support is included in $(0, \infty)$. Unfortunately this contribution has been unnoticed in the literature since it has appeared on an unpublished technical report. If $\alpha \in \mathbb{A}_\theta$, from identity (2.8) it is possible to show that

$$\alpha((a,b]) = \frac{1}{\pi} \lim_{\epsilon \downarrow 0} \int_a^b \mathrm{Im}\left(\frac{\mathcal{S}'_\theta[-x + \mathrm{i}\epsilon;\, M(\tilde{\mathscr{D}}_\alpha)]}{\mathcal{S}_\theta[-x + \mathrm{i}\epsilon;\, M(\tilde{\mathscr{D}}_\alpha)]} \right) \mathrm{d}x \tag{2.25}$$

where a and b are continuity points of A. Below we provide three examples for which it is possible to determine explicitly α starting from a pre-specified law for the mean. See [11] for details and the discussion of the first two examples we describe below.

Example 2.5. Suppose that $q(x;\alpha) = B(\mathrm{d}x; a, b)$ with a and b positive and $a + b = \theta$, where, as before, $B(\mathrm{d}x; a, b)$ denotes a beta probability distribution. Moreover, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. In this case it can be seen that

$$\mathcal{S}_\theta[z; M(\tilde{\mathscr{D}}_\alpha)] = z^{-a}(1+z)^{-b}.$$

An application of (2.25) yields that the parameter measure α has support $\{0,1\}$ and it coincides with $\alpha(\mathrm{d}x) = a\,\delta_0(\mathrm{d}x) + b\,\delta_1(\mathrm{d}x)$. □

Example 2.6. Another interesting case is that of $M(\tilde{\mathscr{D}}_\alpha)$ beta distributed with parameters $\theta + 1/2$ and $\theta + 1/2$. One can check that

$$\mathcal{S}_\theta[z; M(\tilde{\mathscr{D}}_\alpha)] = z^{-\theta} 2^{2\theta} \left(1 + \sqrt{1 + z^{-1}}\right)^{-2\theta}$$

and resorting again to (2.25) one finds out that $\alpha = \theta P_0$, where P_0 is a beta probability distribution with parameters $(1/2, 1/2)$. □

Example 2.7. Suppose $q(x;\alpha) = \mathbb{1}_{(0,1)}(x)$ so that

$$\mathcal{S}_1[z; M(\tilde{\mathscr{D}}_\alpha)] = \log(1+z) - \log(z).$$

This implies that $\mathcal{S}'_1[z; M(\tilde{\mathscr{D}}_\alpha)] = -z^{-1}(1-z)^{-1}$ and for any $x \in (0,1)$

$$\frac{\mathcal{S}'_1[-x + \mathrm{i}\epsilon; M(\tilde{\mathscr{D}}_\alpha)]}{\mathcal{S}_1[-x + \mathrm{i}\epsilon; M(\tilde{\mathscr{D}}_\alpha)]} = \frac{x(1-x) + \epsilon^2 - \mathrm{i}\epsilon(1-2x)}{[x(1-x)+\epsilon^2]^2 + \epsilon^2(1-2x)^2} \times$$
$$\times \frac{\frac{1}{2}\log\frac{(1-x)^2+\epsilon^2}{x^2+\epsilon^2} + \mathrm{i}\big(\pi - \arctan\frac{\epsilon}{1-x} - \arctan\frac{\epsilon}{x}\big)}{\frac{1}{4}\log^2\frac{(1-x)^2+\epsilon^2}{x^2+\epsilon^2} + \big[\pi - \arctan\frac{\epsilon}{1-x} - \arctan\frac{\epsilon}{x}\big]^2}.$$

From the previous expression one easily deduces that

$$\lim_{\epsilon \downarrow 0} \mathrm{Im} \frac{\mathcal{S}'_1[-x + \mathrm{i}\epsilon; M(\tilde{\mathscr{D}}_\alpha)]}{\mathcal{S}_1[-x + \mathrm{i}\epsilon; M(\tilde{\mathscr{D}}_\alpha)]} = \frac{x^{-1}(1-x)^{-1}}{\pi^2 + \log^2\frac{1-x}{x}}. \tag{2.26}$$

Hence, the parameter measure α, corresponding to a random Dirichlet mean $M(\tilde{\mathscr{D}}_\alpha)$ having a uniform distribution on $(0,1)$, admits density (2.26) (w.r.t. the Lebesgue measure on $\mathbb{R}$). As noted in [75, 76], the cumulative distribution function A corresponding to the density function in (2.26) also appears as the limiting curve of the sequence of fractional parts of the roots of the derivative of the polynomial $p_n(x) = (x-n)_{n+1}$. In other words, letting $p_n'(x) = n\prod_{k=0}^{n-1}(x-k-\alpha_{n,k})$ one has

$$\lim_{n\to\infty} \alpha_{n,[xn]} = \frac{1}{\pi}\operatorname{arccot}\left(\frac{1}{\pi}\log\frac{1-x}{x}\right) = \int_0^x \frac{s^{-1}(1-s)^{-1}}{\pi^2+\log^2\frac{1-s}{s}}\,\mathrm{d}s$$

for any $x\in(0,1)$, where $[y]$ stands for the integer part of y. Hence, the limiting curve obtained from the sequence of fractional parts of the roots of p_n' coincides with the Markov transform of the uniform distribution on $(0,1)$. □

The previous three examples show that a uniform distribution for the mean may correspond to different parameter measures according to the specific total mass θ being chosen. This is not in contradiction with the fact that there is a bijection between $\mathbb{M}_\theta$ and $\mathbb{A}_\theta$, since we let θ vary. Hence, if $\theta=1$ the parameter measure has a density of the form described in (2.26), if $\theta=1/2$ then $\alpha = P_0/2$, where P_0 is a beta distribution with parameters $(1/2,1/2)$ and if $\theta=2$, then α is a discrete measure with point masses equal to 1 at 0 and 1.

The problem of determining α for fixed $Q\in\mathbb{M}_\theta$ has been also considered in combinatorics where random Dirichlet means arise in connection with Young continual diagrams which represent a generalization of Young diagrams for partitions. See [49] and [75, 76] for some detailed and exhaustive accounts. In particular, in [75] one finds a result according to which, given a density q in $\mathbb{M}_\theta$ with $\theta=1$ and support in (a,b), then the corresponding α is such that

$$\alpha((a,x]) = \frac{1}{\pi}\operatorname{arccot}\left(\frac{1}{\pi\, q(x;\alpha)}\,\mathrm{PV}\int_a^b \frac{q(u;\alpha)}{u-x}\,\mathrm{d}u\right). \tag{2.27}$$

Hence, combining the representations obtained by [11] and [75] with $\theta=1$, one has the following new identity

$$\lim_{\epsilon\downarrow 0}\int_a^x \mathrm{Im}\left(\frac{\mathcal{S}_1'[-u+\mathrm{i}\epsilon; M(\tilde{\mathscr{D}}_\alpha)]}{\mathcal{S}_1[-u+\mathrm{i}\epsilon; M(\tilde{\mathscr{D}}_\alpha)]}\right)\mathrm{d}u = \operatorname{arccot}\left(\frac{1}{\pi\, q(x;\alpha)}\,\mathrm{PV}\int_a^b \frac{q(u;\alpha)}{u-x}\,\mathrm{d}u\right) \tag{2.28}$$

which does not seem immediate to us. Although not concerning random Dirichlet means, we also point out an allied contribution. In [34] a constructive method, involving the sequential generation of an array of barycenters, is developed in order to define random probability measures with prescribed distribution of their means.

2.4. More distributional properties of random Dirichlet means

In the previous Section we have provided a description of the probability distribution of $M(\tilde{\mathscr{D}}_\alpha)$ in cases where the expression of the density $q(\,\cdot\,;\alpha)$ admits a simple expression. In general, even though expressed in closed form, the density $q(\,\cdot\,;\alpha)$ might turn out to be difficult to handle and it is not helpful in order to determine some characteristics of the probability distribution of $M(\tilde{\mathscr{D}}_\alpha)$ such as its moments or its shape. Nonetheless, one can still devise some method to analyze distributional properties of $M(\tilde{\mathscr{D}}_\alpha)$ even if its density function looks cumbersome.

As for the determination of moments, some early contributions can be found in [87, 88]. In these papers the evaluation of $\mathbb{E}\big[(M(\tilde{\mathscr{D}}_\alpha))^n\big]$ is based on combinatorial arguments and follows from the relation

$$\mathbb{E}\left[\prod_{i=1}^{n}\int_{\mathbb{X}} g_i(x)\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x)\right] = \frac{n!}{(\theta)_n}\sum_{k=1}^{n}\sum_{(m_1,\ldots,m_n)\in Z(n,k)}\frac{1}{\prod_{i=1}^{n} i^{m_i}(m_i!)}\times \\ \times\int_{\mathbb{X}^k} g(\boldsymbol{x})\prod_{i=1}^{n}\prod_{j=1}^{m_i}\alpha(\mathrm{d}x_{i,j}) \quad (2.29)$$

where $g:\mathbb{X}^n\to\mathbb{R}$ is a measurable and symmetric function, $\boldsymbol{x}$ is a vector whose generic entry $x_{i,j}$ is repeated i times for any $j=1,\ldots,m_i$ and the set $Z(n,k)$ is defined in (2.12). If one chooses $g(\boldsymbol{x})\equiv\prod_{i=1}^{n}(\prod_{j=1}^{m_i}x_{i,j})^i$, the n–th moment of the random Dirichlet mean is obtained. This can be summarized in the following statement

Theorem 2.5. *Let $\alpha\in\mathbb{A}_\theta$ be such that $\int_{\mathbb{R}}|x|^n\,\alpha(\mathrm{d}x)<\infty$ and set $r_{i,\alpha}=\int_{\mathbb{R}}x^i\,\alpha(\mathrm{d}x)$. Then*

$$\mathbb{E}\left[(M(\tilde{\mathscr{D}}_\alpha))^n\right] = \frac{n!}{(\theta)_n}\sum_{k=1}^{n}\sum_{(*)}\frac{1}{\prod_{i=1}^{n} i^{m_i}(m_i!)}\prod_{i=1}^{n-k+1} r_{i,\alpha}^{m_i} \\ = \frac{1}{(\theta)_n}\sum_{k=1}^{n} B_{n,k}(r_{1,\alpha},2!r_{2,\alpha},\ldots,(n-k+1)!r_{n-k+1,\alpha}) \quad (2.30)$$

where

$$B_{n,k}(z_1,\ldots,z_{n-k+1}) = \sum_{(m_1,\ldots,m_n)\in Z(n,k)}\frac{n!}{m_1!\cdots m_{n-k+1}!}\prod_{i=1}^{n-k+1}\frac{z_i^{m_i}}{(i!)^{m_i}}$$

is, for any $k\le n$, the partial exponential Bell polynomial.

The result just stated has been independently obtained in a few recent papers. For example, in [15] the expression in (2.30) is deduced as a corollary of the Ewens sampling formula which describes the partition of the first n elements of an infinite sequence of exchangeable random variables directed by a Dirichlet

process. On the other hand, in [71] the connection between random Dirichlet means and multiple hypergeometric functions, as already described in the Section 2.2, is exploited. Indeed, under the assumption that the support of α is a bounded subset of $\mathbb{R}^+$, a power series representation for the moment generating function is derived from which the n–th moment is identified by

$$\mathbb{E}\left[(M(\tilde{\mathscr{D}}_\alpha))^n\right] = \frac{1}{(\theta)_n}\,\mathbb{E}\left[\left(\int x\,\tilde{\gamma}(\mathrm{d}x)\right)^n\right].$$

The expectation on the right–hand side of the previous equality is evaluated by means of the Faà di Bruno formula and this explains the appearance of the partial exponential Bell polynomial $B_{n,k}$ in (2.30). A useful account on the Faà di Bruno formula, together with applications to various examples related to the material presented in this paper, is provided by [13].

Recursive formulae for the moments of random Dirichlet means can be found in [7] and [36]. They have been obtained without conditions on the support of α and under the assumption that $\int |x|^n\,\tilde{\mathscr{D}}_\alpha(\mathrm{d}x)$ is finite. In [7] the starting point is the identity (2.8), whereas the proof in [36] makes use of the fact that the Dirichlet process is the unique solution of the stochastic equation

$$\tilde{\mathscr{P}} \stackrel{d}{=} w\,\delta_{X_1} + (1-w)\,\tilde{\mathscr{P}} \tag{2.31}$$

where $\tilde{\mathscr{P}}$ is a random probability measure on $\mathbb{R}$, the random variable X_1 has probability distribution α/θ, w is a beta distributed random variable with parameters $(1,\theta)$, X_1 and w are independent and the random probability $\tilde{\mathscr{P}}$ on the right–hand side is independent from (w, X_1). See [78]. From this one deduces $M(\tilde{\mathscr{D}}_\alpha) \stackrel{d}{=} wX_1 + (1-w)M(\tilde{\mathscr{D}}_\alpha)$ and, because of independence,

$$\mathbb{E}\left[(M(\tilde{\mathscr{D}}_\alpha))^n\right] = \frac{n!}{\Gamma(\theta+n+1)}\sum_{j=0}^{n}\frac{\Gamma(\theta+j)}{j!}\,\mathbb{E}\left[(M(\tilde{\mathscr{D}}_\alpha))^j\right]\,r_{n-j,\alpha}$$

where we recall that $r_{i,\alpha} = \int_{\mathbb{R}} x^i\alpha(\mathrm{d}x)$ for any i. Consequently, if $\int |x|^n\,\alpha(\mathrm{d}x) < \infty$, one finds out that

$$\mathbb{E}\left[(M(\tilde{\mathscr{D}}_\alpha))^n\right] = (n-1)!\sum_{j=0}^{n-1}\frac{r_{n-j,\alpha}}{j!\,(\theta+j)_{n-j}}\,\mathbb{E}\left[(M(\tilde{\mathscr{D}}_\alpha))^j\right]. \tag{2.32}$$

See [36] for details.

Another interesting distributional property that has been investigated is the connection between the symmetry of $q(\,\cdot\,;\alpha)$ and the symmetry of the corresponding $\alpha \in \mathbb{A}_\theta$. If the parameter measure $\alpha \in \mathbb{A}_\theta$ is symmetric about $c \in \mathbb{R}$, then $M(\tilde{\mathscr{D}}_\alpha)$ has a symmetric distribution about c. Such a result has been independently achieved [87] and [33]. In [87] the proof is easily deduced from the representation of the Dirichlet process as provided in (2.3). Indeed, $\sum_{j\geq 1}\tilde{p}_j\delta_{Y_j} \stackrel{d}{=} \sum_{j\geq 1}\tilde{p}_j\delta_{2c-Y_j}$ because of symmetry of α. Hence

$$M(\tilde{\mathscr{D}}_\alpha) = \sum_{j\geq 1} Y_j\,\tilde{p}_j \stackrel{d}{=} \sum_{j\geq 1}(2c-Y_j)\,\tilde{p}_j = 2c - M(\tilde{\mathscr{D}}_\alpha). \tag{2.33}$$

In [33] a proof is given under the assumption that the symmetry center is $c = 0$. The argument relies on the relation between $M(\tilde{\mathscr{D}}_\alpha)$ and a linear functional of the corresponding gamma process recalled in (2.22). Hence, the evaluation of the characteristic function of the linear functional of $\tilde{\gamma}$ allows one to show that $\int (x-y)\,\tilde{\gamma}(\mathrm{d}x) \stackrel{d}{=} -\int (-x-y)\,\tilde{\gamma}(\mathrm{d}x)$ thus yielding $\mathbb{P}\big[M(\tilde{\mathscr{D}}_\alpha) \le y\big] = \mathbb{P}\big[M(\tilde{\mathscr{D}}_\alpha) \ge -y\big]$. The authors also hint at (2.33) as an alternative proof of the result. See also [36] where similar arguments are used. In [72], the authors obtain the representation $\mathbb{P}[M(\tilde{\mathscr{D}}_\alpha) \le y] = \int_0^\infty h(x,y)\mathrm{d}y$ in terms of a function $h : \mathbb{R}^2 \to \mathbb{R}^+$ such that, if α is symmetric about $c = 0$, then $h(x,y) = h(x,-y)$ for any $y \in \mathbb{R}$ and this obviously implies symmetry of $q(\,\cdot\,;\alpha)$. Finally, in [59] symmetry is proved by showing that the characteristic function of $M(\tilde{\mathscr{D}}_\alpha)$ is real.

Recently, in [42] it has been proved that for any $\theta > 0$ and $\sigma \in (0,1)$, one has

$$w_{\theta\sigma,\theta(1-\sigma)}\, M(\tilde{\mathscr{D}}_\alpha) \stackrel{d}{=} M(\tilde{\mathscr{D}}_{\alpha^*}) \tag{2.34}$$

where $\alpha \in \mathbb{A}_{\theta\sigma}$, $\alpha^* \in \mathbb{A}_\theta$ and $w_{\theta\sigma,\theta(1-\sigma)}$ is a beta random variable with parameters $(\theta\sigma, \theta(1-\sigma))$ that is independent from $M(\tilde{\mathscr{D}}_\alpha)$. The distributional identity in (2.34) is useful when the total mass of the Dirichlet process is less than one, which is the most difficult case to handle. Indeed, if one fixes $\theta = 1$ one has that the random mean of a Dirichlet process with parameter $\sigma \in (0,1)$, rescaled with an independent beta random variable, has the same distribution of the mean of a Dirichlet process whose baseline measure has total mass 1. And the latter can be determined via the results described in Section 2.2. Finally, distributional connections between random Dirichlet means and generalized gamma convolutions have been deeply investigated in [47].

2.5. Numerical results

In the previous Sections we have been dealing with exact results about random Dirichlet means. However, as already highlighted at the beginning of Section 2.4, the form of the density $q(\,\cdot\,;\alpha)$ can be quite complicated to handle for practical purposes. Hence, it is also of interest to determine suitable approximations of the law of $M(\tilde{\mathscr{D}}_\alpha)$. There have been various proposals in the literature which we concisely recall.

An interesting approach we would like to account for is based on a remarkable result obtained in [25]. Inspired by the distributional identity (2.31), they define a recursive equation

$$\tilde{\mathscr{P}}_n = w_n\, \delta_{X_n} + (1 - w_n)\, \tilde{\mathscr{P}}_{n-1} \qquad n \ge 1 \tag{2.35}$$

where $\tilde{\mathscr{P}}_0$ is an arbitrary probability measure on $(\mathbb{X}, \mathscr{X})$, the random elements $\tilde{\mathscr{P}}_{n-1}$, w_n and X_n are mutually independent and $\{(w_n, X_n)\}_{n=1}^\infty$ forms a sequence of i.i.d. random variables with the same distribution as (w, X_1) in (2.31). It is apparent that (2.35) defines a Markov chain taking values in the space of probability measures on $(\mathbb{X}, \mathscr{X})$. Theorem 1 in [25] states that there is a unique invariant measure for the sequence of random probability measures $\{\tilde{\mathscr{P}}_n\}_{n\ge 0}$

on $(\mathbb{X}, \mathscr{X})$ defined by (2.35) and it coincides with $\tilde{\mathscr{D}}_\alpha$. This finding leads to study properties of the Markov chain $\{M_n\}_{n\geq 0}$ with $M_n = \int x \tilde{\mathscr{P}}_n(dx)$ and it can be shown that, for $\alpha \in \mathbb{A}_\theta$, $\{M_n\}_{n\geq 0}$ is a Harris ergodic Markov chain whose unique invariant distribution coincides with $Q(\cdot;\alpha)$. See Theorem 1 in [21] for a multidimensional extension of these results. Consequently, one may try to devise a Markov chain Monte Carlo (MCMC) algorithm to draw a sample having approximately probability distribution $Q(\,\cdot\,;\alpha)$. This goal has been pursued in [29] where it is shown that the chain $\{M_n\}_{n\geq 0}$ is geometrically ergodic if $\int |x|\, \alpha(\mathrm{d}x) < \infty$. Such an integrability condition has been subsequently relaxed in [48]. If it is further assumed that supp(α) is bounded, one has the stronger property of uniform ergodicity. See Theorem 4 in [29]. Interestingly, the authors are also able to determine upper bounds on the rate of convergence thus obtaining an evaluation of the total variation distance between $Q(\,\cdot\,;\alpha)$ and the law of the approximating mean M_n. As a subsequent development along these lines, [30] formulate an algorithm which allows for perfect sampling from $Q(\,\cdot\,;\alpha)$. Recently, a generalization of the Markov chain in (2.35) has been introduced and investigated in [24].

Another simulation–based method has been proposed by [62]. The main idea relies on truncating the series representation (2.3) of $\tilde{\mathscr{D}}_\alpha$ at some random point N in such a way that the distance between $\tilde{\mathscr{D}}_\alpha$ and its truncated version is (almost surely) below a certain threshold. Define the random variable $N_\epsilon = \inf\{m \geq 1 : \sum_{i=1}^m \tilde{p}_i < 1-\epsilon\}$ and let $\tilde{r}_\epsilon = \sum_{i\geq N_\epsilon+1} \tilde{p}_i$. Hence

$$\tilde{\mathscr{P}}_\epsilon = \sum_{i=1}^{N_\epsilon} \tilde{p}_i\, \delta_{Y_i} + \tilde{r}_\epsilon\, \delta_{Y_0}$$

with Y_0 a random variable with probability distribution α/θ and independent from the sequences of stick–breaking weights $\{V_n\}_{n\geq 1}$ and from the random variables $\{Y_n\}_{n\geq 1}$. In [62] $\tilde{\mathscr{P}}_\epsilon$ is termed ϵ–Dirichlet random probability. The authors show that if d_p denotes the Prokhorov distance between probability distributions then $d_p(\tilde{p}_\epsilon, \tilde{\mathscr{D}}_\alpha) \leq \epsilon$, almost surely. Moreover, one can check that N_ϵ has a shifted Poisson distribution with parameter $-\theta \log \epsilon$ so that it is easy to sample from $\tilde{\mathscr{P}}_\epsilon$. In some cases it is possible to show that the closeness, with respect to the Prokhorov distance, between $\tilde{\mathscr{P}}_\epsilon$ and $\tilde{\mathscr{D}}_\alpha$ induces closeness between the probability distributions of $\int x\, \tilde{\mathscr{P}}_\epsilon(\mathrm{d}x)$ and of $M(\tilde{\mathscr{D}}_\alpha)$.

As far as numerical methods for approximating $Q(\,\cdot\,;\alpha)$ are concerned, a first application can be found in [83] and it is based on a numerical inversion of the Laplace functional transform of the gamma process $\tilde{\gamma}$. Another possibility consists in approximating the cumulative distribution function $\mathbb{F}(\,\cdot\,)$. Since the expression in (2.24) can be evaluated, at least numerically, when the parameter measure has finite support, in general, one might approximate $\mathbb{F}(\,\cdot\,)$ with the cumulative distribution function of $\int x \mathrm{d}\tilde{\mathscr{D}}_{\alpha_n}$, in symbols $\mathbb{F}_n(\,\cdot\,)$, where α_n is finite measure supported by $\{x_{1,n}, \ldots, x_{k_n,n}\}$. Given a prescribed error of approximation $\epsilon > 0$, in [72] one can find a method for determining constants $U < L$ and

points $x_{1,n}, \ldots, x_{k_n,n}$ in $(U, L]$ such that

$$\sup_{y \in (U,L]} |\mathbb{F}(y) - \mathbb{F}_n(y)| < \epsilon$$

When the support of α is bounded, U and L coincide with its upper and lower bound, respectively. One can, then, evaluate a cumulative distribution function which approximates, in the uniform metric, the actual cumulative distribution function $\mathbb{F}(\,\cdot\,)$ on a sufficiently large interval of the real line.

3. Means of more general processes

The Dirichlet process has been the starting point for the development of Bayesian nonparametric methods and it still is one of the most widely used nonparametric priors. Nonetheless some of the properties featured by the Dirichlet process may represent drawbacks in various inferential contexts. This has stimulated a strong effort towards the definition of classes of nonparametric priors which extend the Dirichlet process. Accordingly, one might try to generalize results known for the Dirichlet case to random means of these more general random probability measures. In the present Section we provide an overview of these contributions which incidentally have interesting connections with other areas of research such as the study of the excursions of stochastic processes, the investigation of fugacity measures in physics and the analysis of option pricing in mathematical finance. Section 3.1 illustrates a few recent distributional results involving linear functionals of the two–parameter Poisson–Dirichlet process. Means of normalized random measures with independent increments are the focus of Section 3.2, whereas in Section 3.3 we deal with means of neutral to the right processes.

3.1. Two parameter Poisson-Dirichlet process

An important extension of the Dirichlet process has been proposed by Pitman [65] as a two parameter exchangeable random partition on the set of integers $\mathbb{N}$. It is termed two–parameter Poisson–Dirichlet process since it is seen as a generalization of the one–parameter Poisson–Dirichlet process introduced by Kingman in [52] which identifies the probability distribution of the random Dirichlet probabilities in the representation (2.3) ranked in an increasing order.

For our purposes, it is convenient to define the two–parameter Poisson–Dirichlet process by means of a rescaling of the probability distribution of a σ–stable CRM $\tilde{\mu}_\sigma$, with $\sigma \in (0, 1)$ and base measure P_0 being a probability measure on $(\mathbb{X}, \mathscr{X})$. This means that $\tilde{\mu}_\sigma$ is a random element taking values in the space $(\mathcal{M}_\mathbb{X}, \mathscr{M}_\mathbb{X})$ of boundedly finite measure on $(\mathbb{X}, \mathscr{X})$ such that for any measurable function $f : \mathbb{X} \to \mathbb{R}$ satisfying $\int |f| \, \mathrm{d}\tilde{\mu}_\sigma < \infty$ (almost surely), one has $\mathbb{E}\left[\exp\{-\int f \, \mathrm{d}\tilde{\mu}_\sigma\}\right] = \exp\{-\int f^\sigma \, \mathrm{d}\alpha\}$. If $\mathbb{P}_{(\sigma,0)}$ denotes the probability distribution of $\tilde{\mu}_\sigma$, introduce $\mathbb{P}_{(\sigma,\theta)}$ on $\mathcal{M}_\mathbb{X}$ such that $\mathbb{P}_{(\sigma,\theta)} \ll \mathbb{P}_{(\sigma,0)}$ and

$$\frac{\mathrm{d}\mathbb{P}_{(\sigma,\theta)}}{\mathrm{d}\mathbb{P}_{(\sigma,0)}}(\mu) = [\mu(\mathbb{X})]^{-\theta} \tag{3.1}$$

where $\theta > -\sigma$. The random measure with probability distribution $\mathbb{P}_{(\sigma,\theta)}$ is denoted by $\tilde{\mu}_{(\sigma,\theta)}$ and the random probability measure

$$\tilde{\mathscr{P}}_{(\sigma,\theta)} \stackrel{d}{=} \frac{\tilde{\mu}_{(\sigma,\theta)}(\,\cdot\,)}{\tilde{\mu}_{(\sigma,\theta)}(\mathbb{X})} \tag{3.2}$$

is termed *two–parameter Poisson–Dirichlet process* and we will also use the notation $\mathrm{PD}(\sigma,\theta)$ when referring to it. Note that $\mathbb{E}\big[\tilde{\mathscr{P}}_{(\sigma,\theta)}(B)\big] = P_0(B)$ for any B in $\mathscr{X}$. Like the Dirichlet process, one can also provide a stick–breaking construction of $\tilde{\mathscr{P}}_{(\sigma,\theta)}$. Indeed, if $\{W_n\}_{n\geq 1}$ is a sequence of independent random variables with $W_n \sim \mathrm{beta}(1-\sigma, \theta+n\sigma)$, and if one defines $\tilde{p}_1 = W_1$, $\tilde{p}_j = W_j \prod_{i=1}^{j-1}(1-W_i)$ for $j \geq 2$, then

$$\tilde{\mathscr{P}}_{(\sigma,\theta)} \stackrel{d}{=} \sum_{j\geq 1} \tilde{p}_j\, \delta_{Y_j} \tag{3.3}$$

where $\{Y_n\}_{n\geq 1}$ is a sequence of i.i.d. random variables with probability distribution P_0. This construction has become very popular in Bayesian Nonparametrics practice due to the fact it allows a simple scheme for simulating $\tilde{\mathscr{P}}_{(\sigma,\theta)}$. The Dirichlet process can be seen as a two–parameter Poisson–Dirichlet process: if α has total mass $\theta > 0$, then $\tilde{\mathscr{D}}_\alpha \stackrel{d}{=} \tilde{\mathscr{P}}_{(0,\theta)}$.

In order to evaluate the probability distribution of the random mean $M(\tilde{\mathscr{P}}_{(\sigma,\theta)}) := \int_{\mathbb{R}} x\, \tilde{\mathscr{P}}_{(\sigma,\theta)}(\mathrm{d}x)$, we consider probability measures P_0 such that

$$\int |x|^\sigma \,\mathrm{d}P_0 < \infty \tag{3.4}$$

since the integrability of $|x|^\sigma$ with respect to P_0 is a necessary and sufficient condition for $\int |x|\, \tilde{\mathscr{P}}_{(\sigma,\theta)}(\mathrm{d}x) < \infty$ with probability one. See Proposition 1 in [73]. Moreover, by virtue of Theorem 2.1 in [44], the probability distribution of $M(\tilde{\mathscr{P}}_{(\sigma,\theta)})$ is absolutely continuous with respect to the Lebesgue measure. We denote the corresponding density function as $q_{(\sigma,\theta)}$. An important characterization of $q_{(\sigma,\theta)}$ has been first provided in [50] where it is shown that

$$\mathcal{S}_\theta\left[z; M(\tilde{\mathscr{P}}_{(\sigma,\theta)})\right] = \left\{\int_{\mathbb{R}} [z+x]^\sigma\, P_0(\mathrm{d}x)\right\}^{-\theta/\sigma} \tag{3.5}$$

for any $z \in \mathbb{C}$ such that $|\arg(z)| < \pi$, for any $\sigma \in (0,1)$ and $\theta > -\sigma$. Alternative proofs of the identity in (3.5) can be found in [84] and in [85].

The generalized Stieltjes transform in (3.5) can be taken as the starting point for the determination of $q_{(\sigma,\theta)}$. Inspired by the work of [10], one can try to invert $\mathcal{S}_\theta$ and evaluate the probability distribution of $M(\tilde{\mathscr{P}}_{(\sigma,\theta)})$. This goal has been achieved in [44]. Before describing the representations one can obtain for the density $q_{(\sigma,\theta)}$ and the corresponding cumulative distribution function, we wish to emphasize that the analysis of $M(\tilde{\mathscr{P}}_{(\sigma,\theta)})$ is of interest also beyond the Bayesian Nonparametrics. For example, it plays an important role in the analysis of the excursions of skew Bessel bridges. To understand this connection, one can follow [69] and let $Y = \{Y_t : t \geq 0\}$ denote a real–valued process such that: (i) the zero

set Z of Y is the range of a σ–stable subordinator and (ii) given $|Y|$, the signs of excursions of Y away from zero are chosen independently of each other to be positive with probability p and negative with probability $1-p$. Examples of this kind of process are: Brownian motion ($\sigma = p = 1/2$); skew Brownian motion ($\sigma = 1/2$ and $0 < p < 1$); symmetrized Bessel process of dimension $2-2\sigma$; skew Bessel process of dimension $2-2\sigma$. Then for any random time T which is a measurable function of $|Y|$

$$\frac{A_T}{T} = \frac{1}{T}\int_0^T \mathbb{1}_{(0,+\infty)}(Y_s)\,\mathrm{d}s \stackrel{d}{=} \int_0^1 \mathbb{1}_{(0,+\infty)}(Y_s)\,\mathrm{d}s = A_1 \tag{3.6}$$

with A_1 representing the time spent positive by Y between $t=0$ and $t=1$. Interestingly, from (4.b') in [1] one has

$$\mathcal{S}_\sigma[z; A_1] = \frac{1}{(1-p)\,z^\sigma + p\,(1+z)^\sigma} = \mathcal{S}_\sigma\left[z; \tilde{\mathscr{P}}_{(\sigma,\sigma)}(C)\right]$$

with $C \in \mathscr{X}$ being such that $P_0(C) = p$. This implies the remarkable distributional identity $A_1 \stackrel{d}{=} \tilde{\mathscr{P}}_{(\sigma,\sigma)}(C)$. Let, now, $\tilde{\mathscr{P}}'_{(\sigma,\sigma)}$ be a PD(σ,σ) process with $\mathbb{E}\big[\tilde{\mathscr{P}}_{(\sigma,\sigma)}(B)\big] = p\delta_1(B) + (1-p)\,\delta_0(B)$ for any $B \in \mathscr{X}$. The determination of the probability distribution of

$$\tilde{\mathscr{P}}_{(\sigma,\sigma)}(C) = \int x\tilde{\mathscr{P}}'_{(\sigma,\sigma)}(\mathrm{d}x)$$

coincides with the problem of determining the probability distribution of A_1. P. Lévy in [55] has shown that the density function of A_1 is uniform in the case where $p = \sigma = 1/2$. The case where $\theta = 0$ is also interesting since one has that the probability distribution of $\tilde{\mathscr{P}}_{(\sigma,0)}(C)$ coincides with the generalized arcsine law determined in [54]. Another case for which the probability distribution of $\tilde{\mathscr{P}}_{(\sigma,\theta)}$ is known corresponds to $\sigma = 1/2$, $\theta > 0$ and $p \in (0,1)$ (see [5]).

In [44] one can find a general result, which holds for a large class of parameter measures P_0 including the one with point masses in 0 and 1 as above.

Theorem 3.1. *Let P_0 be a probability measure on $\mathbb{R}^+$ such that $\int_{\mathbb{R}^+} x^\sigma\,P_0(\mathrm{d}x) < \infty$ and let*

$$\Delta_{(\sigma,\theta)}(t;\lambda) := \frac{1}{2\pi\mathrm{i}}\lim_{\epsilon\downarrow 0}\left\{\mathcal{S}_\lambda\left[-t-\mathrm{i}\epsilon;\, M(\tilde{\mathscr{P}}_{(\sigma,\theta)})\right] - \mathcal{S}_\lambda\left[-t+\mathrm{i}\epsilon;\, M(\tilde{\mathscr{P}}_{(\sigma,\theta)})\right]\right\}.$$

for any $\lambda > 0$, $\sigma \in (0,1)$ and $\theta > 0$. Then

$$q_{(\sigma,\theta)}(y) = \theta\int_0^y (y-t)^{\theta-1}\,\Delta_{(\sigma,\theta)}(t;\,\theta+1)\,\mathrm{d}t \tag{3.7}$$

for any y in the convex hull of the support of P_0.

If $\theta > 1$ it is possible to integrate by parts in (3.7) thus obtaining

$$q_{(\sigma,\theta)}(y) = (\theta-1)\int_0^y (y-t)^{\theta-2}\,\Delta_{(\sigma,\theta)}(t;\theta)\,\mathrm{d}t. \tag{3.8}$$

For the special case where $\theta = 1$, one uses the Perron–Stieltjes inversion formula which yields $q_{(\sigma,1)}(\cdot) = \Delta_{(\sigma,1)}(\cdot\,;1)$. Here below we illustrate a few examples where the above result is applied.

Example 3.1. When $\theta = 0$, then $\tilde{\mathscr{P}}_{(\sigma,0)}$ reduces to the normalized σ–stable subordinator [52]. Let $\mathscr{C}(P_0)$ denote the convex hull of the support of P_0 and set $\mathbb{S}_\sigma(P_0) = \{y \in \mathbb{R}^+ : \int_{\mathbb{R}^+} |y-t|^\sigma\, P_0(\mathrm{d}t) < \infty\} \cap \mathscr{C}(P_0)$. Moreover, set

$$\mathscr{A}_d^+[t, P_0] := \int_t^\infty (x-t)^d\, P_0(\mathrm{d}x) \quad \text{and} \quad \mathscr{A}_d[t, P_0] := \int_0^t (t-x)^d\, P_0(\mathrm{d}x).$$

If the Lebesgue measure of the set $\mathbb{R}^+ \setminus \mathbb{S}_\sigma(P_0)$ is zero, then one has

$$q_{(\sigma,0)}(y) = \frac{\sin(\sigma\pi)}{\pi}\, \frac{\mathscr{A}_\sigma^+[y;P_0]\,\mathscr{A}_{\sigma-1}[y;P_0] + \mathscr{A}_{\sigma-1}^+[y;P_0]\,\mathscr{A}_\sigma[y;P_0]}{\{\mathscr{A}_\sigma^+[y;P_0]\}^2 + 2\cos(\sigma\pi)\mathscr{A}_\sigma^+[y;P_0]\mathscr{A}_\sigma[y;P_0] + \{\mathscr{A}_\sigma[y;P_0]\}^2}$$

for any $y \in \mathbb{S}_\sigma(P_0)$. An expression for the cumulative distribution function $\int_0^y q_{(\sigma,0)}(x)\,\mathrm{d}x$ can be found in [73]. See the next section for more details on this point. If $P_0 = p\delta_1 + \bar{p}\delta_0$, with $\bar{p} := 1-p$, then

$$q_{(\sigma,0)}(y) = \frac{\sin(\sigma\pi)}{\pi}\, \frac{p\bar{p}y^{\sigma-1}(1-y)^{\sigma-1}\,\mathbb{1}_{(0,1)}(y)}{p^2(1-y)^{2\sigma} + 2\cos(\sigma\pi)p\bar{p}y^\sigma(1-y)^\sigma + \bar{p}^2y^{2\sigma}}$$

which corresponds with Lamperti's generalized arcsine law. See [54]. □

Example 3.2. Introduce the following quantities

$$\gamma_d(t) = \cos(d\pi)\mathscr{A}_d[t;P_0] + \mathscr{A}_d^+[t;P_0], \qquad \zeta_d(t) = \sin(d\pi)\mathscr{A}_d[t;P_0]$$

and

$$\rho_{(\sigma,\theta)}(t) = \frac{\theta}{\sigma}\,\arctan\frac{\zeta_\sigma(t)}{\gamma_\sigma(t)} + \frac{\pi\,\theta}{\sigma}\,\mathbb{1}_{\Gamma_\sigma}(t)$$

where $\Gamma_\sigma := \{t \in \mathbb{R}^+ : \gamma_\sigma(t) < 0\}$. Hence, if we set $\theta < 1$, we can resort to Theorem 5.3 in [44] which yields

$$\Delta_{(\sigma,\theta)}(t;\theta+1) = \frac{\gamma_{\sigma-1}(t)\,\sin\left(\rho_{(\sigma,\theta+\sigma)}(t)\right) - \zeta_{\sigma-1}(t)\,\cos\left(\rho_{(\sigma,\theta+\sigma)}(t)\right)}{\pi\,[\zeta_\sigma^2(t)+\gamma_\sigma^2(t)]^{(\theta+\sigma)/(2\sigma)}}.$$

Consider, again, $P_0 = p\delta_1 + (1-p)\delta_0$ and set $\sigma = \theta$. Recall that $\bar{p} = 1-p$ and one, then, has

$$\begin{aligned}\gamma_d(t) =& \cos(d\pi)\{p(t-1)^d\mathbb{1}_{[1,\infty)}(t) + \bar{p}t^d\mathbb{1}_{(0,1)}(t)\} + \bar{p}|t|^d\mathbb{1}_{(-\infty,0]}(t)\\ &+ p(1-t)^d\mathbb{1}_{(-\infty,1]}(t)\end{aligned}$$

$$\zeta_d(t) = \sin(d\pi)\,\{p(t-1)^d\mathbb{1}_{[1,\infty)}(t) + \bar{p}t^d\mathbb{1}_{(0,1)}(t)\}$$

$\sin(\rho_{\sigma,2\sigma}(t)) = \sin(2\arctan(\zeta_\sigma(t)/\gamma_\sigma(t)))$ and $\cos(\rho_{\sigma,2\sigma}(t)) = \cos(2\arctan(\zeta_\sigma(t)/\gamma_\sigma(t)))$. Hence, for any $t \in (0,1)$,

$$\Delta_{(\sigma,\sigma)}(t;\sigma+1) = \frac{1}{\pi}\,\frac{2\gamma_{\sigma-1}(t)\zeta_\sigma(t)\gamma_\sigma(t) - \gamma_\sigma^2(t)\zeta_{\sigma-1}(t) + \zeta_{\sigma-1}(t)\zeta_\sigma^2(t)}{\left[\bar{p}^2t^{2\sigma} + p^2(1-t)^{2\sigma} + 2\cos(\sigma\pi)p\bar{p}t^\sigma(1-t)^\sigma\right]^2}$$

and a density function for the time A_1 spent positive by a skew Bessel bridge, as defined in (3.6), is

$$q_{(\sigma,\sigma)}(y) = \frac{\sigma}{\pi} \int_0^y (y-t)^{\sigma-1} \frac{2\gamma_{\sigma-1}(t)\zeta_\sigma(t)\gamma_\sigma(t) - \gamma_\sigma^2(t)\zeta_{\sigma-1}(t) + \zeta_{\sigma-1}(t)\zeta_\sigma^2(t)}{[\bar{p}^2 t^{2\sigma} + p^2(1-t)^{2\sigma} + 2\cos(\sigma\pi)p\bar{p}t^\sigma(1-t)^\sigma]^2}\,\mathrm{d}t$$

for any y in $(0,1)$. See Corollary 6.2 in [44] for details. It is worth mentioning that in [90] one can find an investigation on the behavior of $q_{(\sigma,\sigma)}$ in a neighbourhood of $y=0$. Choosing $\sigma = 1/2$ leads to some useful simplification since one has

$$\Delta_{(1/2,1/2)}(t;3/2) = \frac{\bar{p}}{\pi} \frac{p^2(1+t) - \bar{p}^2 t}{\sqrt{t}\,\{\bar{p}^2 t + p^2(1-t)\}^2}$$

for any t in $(0,1)$. If it is further assumed that $p = 1/2$, then

$$\Delta_{(1/2,1/2)}(t;3/2) = \frac{2}{\pi}\, t^{-1/2}$$

and from (3.7) it follows that $q_{1/2,1/2}(y) = \mathbb{1}_{(0,1)}(y)$ which corresponds to the result obtained by [55]. □

The two examples we have been considering assume that P_0 is a Bernoulli probability distribution. One can obviously consider more general probability distributions given the kind of result provided by Theorem 3.1. For additional explicit formulae check Section 6 in [44].

Up to now we have mentioned an approach to the determination of $q_{(\sigma,\theta)}$ which makes use of an inversion formula for its Stieltjes transform. One can, however, rely on other tools to recover $q_{(\sigma,\theta)}$. For example, from (3.3) one notes that $\tilde{\mathscr{P}}_{(\sigma,\theta)} \stackrel{d}{=} W\,\delta_Y + (1-W)\sum_{j\geq 1}\tilde{p}_j^*\,\delta_{Y_j^*}$, where the $\tilde{p}_j^*$ are determined via a stick–breaking procedure involving a sequence of independent random variables $\{W_n\}_{n\geq 1}$ with W_n being beta–distributed with parameters $(1-\sigma, \theta+\sigma+n\sigma)$. Moreover $W \sim \text{beta}(1-\sigma, \theta+\sigma)$ and $Y, Y_1^*, Y_2^*, \ldots$ are i.i.d. random variables with common probability distribution P_0 on $\mathbb{R}^+$ such that $\int x\mathrm{d}P_0 < \infty$. Accordingly, one has

$$M(\tilde{\mathscr{P}}_{(\sigma,\theta)}) \stackrel{d}{=} W\,Y + (1-W)\,M(\tilde{\mathscr{P}}_{(\sigma,\theta+\sigma)}). \tag{3.9}$$

This turns out to be useful when $\theta = 1-\sigma$, in which case (3.9) reduces to

$$M(\tilde{\mathscr{P}}_{(\sigma,1-\sigma)}) \stackrel{d}{=} W\,Y + (1-W)\,M(\tilde{\mathscr{P}}_{(\sigma,1)})$$

and $W \sim \text{beta}(1-\sigma, 1)$. From Theorem 3.1 one obtains

$$q_{(\sigma,1)}(y) = \Delta_{(\sigma,1)}(y;1) = \frac{\sin\left(\frac{1}{\sigma}\arctan\frac{\zeta_\sigma(y)}{\gamma_\sigma(y)} + \frac{\pi}{\sigma}\mathbb{1}_{\Gamma_\sigma}(y)\right)}{\pi\,[\zeta_\sigma^2(y) + \gamma_\sigma^2(y)]^{\frac{1}{2\sigma}}} \tag{3.10}$$

for any y in $\mathbb{S}_\sigma(P_0)$ and this can be used to recover $q_{(\sigma,1-\sigma)}$.

Example 3.3. Let $P_0 = p\delta_1 + \bar{p}\delta_0$ so that the random mean $M(\tilde{\mathscr{P}}_{(\sigma,1-\sigma)})$ coincides in distribution with a limiting random variable arising in phylogenetic trees as stated in Theorem 20 of [32]. An expression for its probability density function can be deduced as follows

$$q_{(\sigma,1-\sigma)}(y) = \mathbb{1}_{(0,1)}(y)\,(1-\sigma)\int_{\mathbb{R}}\int_0^1 w^{-\sigma}\,q_{(\sigma,1)}\left(\frac{y-wx}{1-w}\right)\,\mathrm{d}w\,P_0(\mathrm{d}x)$$

where $q_{(\sigma,1)}$ is the density function described in (3.10). The expression simplifies considerably when $\sigma = 1/2$, since from (3.10) one would deduce

$$q_{(1/2,1)}(y) = \frac{2p\bar{p}\,\sqrt{y(1-y)}}{\pi\{p^2(1-y)+\bar{p}^2y\}^2}\,\mathbb{1}_{(0,1)}(y)$$

and a density function for $M(\tilde{\mathscr{P}}_{(\sigma,1-\sigma)})$ is

$$q_{(1/2,1/2)}(y) = \frac{p}{2}\int_0^1 \frac{1}{\sqrt{w}}\,q_{(1/2,1)}\left(\frac{y-w}{1-w}\right)\,\mathrm{d}w + \frac{\bar{p}}{2}\int_0^1 \frac{1}{\sqrt{w}}\,q_{(1/2,1)}\left(\frac{y}{1-w}\right)\,\mathrm{d}w$$

for any y in $(0,1)$. Suitable change of variables in the two integrals lead to

$$q_{(1/2,1/2)}(y) = \frac{p(1-y)}{2\pi}\int_0^y \frac{1}{\sqrt{y-w}\,(1-w)^{3/2}}\frac{2p\bar{p}\sqrt{w(1-w)}}{\{p^2-(2p-1)w\}^2}\,\mathrm{d}w\,\mathbb{1}_{(0,1)}(y)$$
$$+\frac{\bar{p}y}{2\pi}\int_y^1 \frac{1}{\sqrt{w-y}\,w^{3/2}}\frac{2p\bar{p}\sqrt{w(1-w)}}{\{p^2-(2p-1)w\}^2}\,\mathrm{d}w\,\mathbb{1}_{(0,1)}(y).$$

If one assumes that $p \in (\sqrt{2}-1, 1)$, so that $(2p-1)/p^2 < 1$, a power series expansion of the above integrands allows to obtain a new completely explicit expression for this case

$$q_{(1/2,1/2)}(y) = \mathbb{1}_{(0,1)}(y)\,\frac{\bar{p}\,y(1-y)}{2p^2}\sum_{n,m\geq 0}\frac{\left(\frac{3}{2}\right)_{n+m}(2)_n}{n!(2)_{n+m}}\left(\frac{(2p-1)y}{p^2}\right)^n y^m +$$
$$+\,\mathbb{1}_{(0,1)}(y)\,\frac{p\bar{p}^2\,y(1-y)}{2(2p^2+2p-1)}\times$$
$$\times\sum_{n,m\geq 0}\frac{\left(\frac{3}{2}\right)_{n+m}(2)_n}{n!(2)_{n+m}}\left(\frac{(2p-1)(1-y)}{2p^2+2p-1}\right)^n(1-y)^m$$
$$= \mathbb{1}_{(0,1)}(y)\,\frac{\bar{p}\,y(1-y)}{2p^2}\,F_1\left(\frac{3}{2},2,2;2;\frac{(2p-1)y}{p^2},y\right) +$$
$$+\,\mathbb{1}_{(0,1)}(y)\,\frac{p\bar{p}^2\,y(1-y)}{2(2p^2+2p-1)}\,F_1\left(\frac{3}{2},2,2;2;\frac{(2p-1)(1-y)}{2p^2+2p-1},1-y\right)$$

where F_1 is the first Appell hypergeometric function. See definition in Section 4 of [23]. □

We finally remark that there is a nice distributional connection between $M(\tilde{\mathscr{P}}_{(\sigma,\theta)})$ and random Dirichlet means examined in Section 2.2. Indeed, if P_0 is a probability distribution on $\mathbb{R}^+$ such that $\int x^\sigma \mathrm{d}P_0 < \infty$ and $P_{(\sigma,0)}$ stands for the probability distribution of $M(\tilde{\mathscr{P}}_{(\sigma,0)})$, with $\mathbb{E}\big[\tilde{\mathscr{P}}_{(\sigma,0)}\big] = P_0$ then

$$M(\tilde{\mathscr{P}}_{(\sigma,\theta)}) \stackrel{d}{=} M(\tilde{\mathscr{D}}_\alpha)$$

where $\alpha = \theta\, P_{(\sigma,0)}$. See Theorem 2.1 in [44]. Hence, one can also try to resort to expressions for the probability distributions described in Section 2.2 to obtain a representation for the density function $q_{(\sigma,\theta)}$.

3.2. NRMI

Another interesting family of random probability measures, which includes the Dirichlet process as a special case, is represented by the class of *normalized random measures with independent increments* (NRMI). These were introduced in [73] by drawing inspiration from Ferguson's [26] construction of the Dirichlet process as normalized gamma process and from Kingman's [52] definition of a normalized stable process. Such random probability measures have shown to be a useful tool for inferential purposes: see, for instance, [45] for a comprehensive Bayesian analysis and [58] for an application to mixture modeling.

A NRMI on a Polish space $\mathbb{X}$ can be defined as

$$\tilde{\mathscr{P}}(\,\cdot\,) = \frac{\tilde{\mu}(\,\cdot\,)}{\tilde{\mu}(\mathbb{X})} \tag{3.11}$$

where $\tilde{\mu}$ is a CRM on $\mathbb{X}$ such that $0 < \tilde{\mu}(\mathbb{X}) < \infty$ almost surely.

Since a CRM is identified by the corresponding Poisson intensity ν, which satisfies (A.3), it is natural to express both the finiteness and positiveness conditions of $\tilde{\mu}(\mathbb{X})$ in terms of ν as defined in (A.5). Finiteness of $\tilde{\mu}(\mathbb{X})$ corresponds to requiring its Laplace exponent $\int_{\mathbb{X}\times\mathbb{R}^+}(1-\mathrm{e}^{-\lambda v})\rho_x(\mathrm{d}v)\alpha(\mathrm{d}x)$ to be finite for any $\lambda \geq 0$; if ν is homogeneous, the previous condition is equivalent to asking for $\theta := \alpha(\mathbb{X}) < \infty$. Positiveness is ensured by requiring the existence of a set A such that $\alpha(A) > 0$ and $\int_A \rho_x(\mathbb{R}^+)\alpha(\mathrm{d}x) = \infty$; if ν is homogeneous, it reduces to $\rho(\mathbb{R}^+) = \infty$, which is equivalent to requiring that the CRM $\tilde{\mu}$ has infinitely many jumps on any bounded set. Such a property is often called *infinite activity* of $\tilde{\mu}$. See [73] for details. In the following NRMIs are termed *homogeneous* or *non–homogeneous* according as to whether they are based on homogeneous or non–homogeneous CRMs. As recalled in (2.1) the Dirichlet process can be seen as a gamma NRMI. Moreover, the two extreme cases of the two–parameter Poisson–Dirichlet process identified by the pairs $(0,\theta)$ and $(\sigma,0)$, with $\theta > 0$ and $\sigma \in (0,1)$, are NRMIs: they coincide with the Dirichlet process and the stable NRMI, respectively, where the latter corresponds to (3.11) with $\tilde{\mu}$ being a stable CRM. On the other hand, the $\tilde{\mathscr{P}}_{(\sigma,\theta)}$ process with other choices of pairs (σ,θ) does not belong to the class of NRMIs, although they are closely connected. See [67].

When studying the distributional properties of a linear functional of $\tilde{\mathscr{P}}$, i.e. $\int_{\mathbb{X}} f(x)\tilde{\mathscr{P}}(\mathrm{d}x)$ with $f : \mathbb{X} \to \mathbb{R}$, one first has to determine conditions which ensure its finiteness. According to Proposition 1 in [73], for any NRMI $\tilde{\mathscr{P}}$ one has that

$$\mathbb{P}\left\{\int_{\mathbb{X}} |f(x)|\tilde{\mathscr{P}}(\mathrm{d}x) < \infty\right\} = 1$$

if and only if

$$\int_{\mathbb{X}\times\mathbb{R}^+} (1 - \mathrm{e}^{-\lambda v\,|f(x)|})\rho_x(\mathrm{d}v)\alpha(\mathrm{d}x) < \infty \quad \text{for every} \quad \lambda > 0. \tag{3.12}$$

Analogously to the Dirichlet and two–parameter Poisson–Dirichlet cases, one may confine the analysis, without loss of generality, to studying a simple mean $M(\tilde{\mathscr{P}}) := \int_{\mathbb{R}} x\,\tilde{\mathscr{P}}(\mathrm{d}x)$. Indeed, letting $\tilde{\mathscr{P}}$ be a NRMI with intensity $\rho_x(\mathrm{d}v)\alpha(\mathrm{d}x)$ and $\alpha_f = \alpha \circ f^{-1}$ denote the image measure of α through f, one has

$$\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{P}}(\mathrm{d}x) \stackrel{d}{=} \int_{\mathbb{R}} x\,\tilde{\mathscr{P}}'(\mathrm{d}x),$$

where $\tilde{\mathscr{P}}'$ is a NRMI with Poisson intensity $\rho_x(\mathrm{d}v)\alpha_f(\mathrm{d}x)$.

We start by briefly discussing two interesting distributional properties already presented for random Dirichlet means, which hold true also for homogeneous NRMIs. The first property is related to the Cauchy distribution: as shown in Example 2.4, $P_0 = \alpha/\theta$ has the same distribution of $M(\tilde{\mathscr{D}}_\alpha)$ if P_0 is Cauchy and, indeed, the same applies to homogeneous NRMI means $M(\tilde{\mathscr{P}})$. The second property is represented by the symmetry of the distribution of the mean discussed at the end of Section 2.4 for Dirichlet means: we have that, if $M(\tilde{\mathscr{P}})$ is a homogeneous NRMI mean, then its distribution is symmetric about $c \in \mathbb{R}$ if P_0 is symmetric about $c \in \mathbb{R}$. These facts, noted in [46], can be shown by simply mimicking known proofs for the Dirichlet case. Indeed, such properties hold for an even larger class of models, namely *species sampling models* (see [66]) which include, besides homogeneous NRMIs, also the two–parameter Poisson–Dirichlet process as noteworthy examples.

Now we turn attention to the problem of the determination of the distribution of a random NRMI mean. In view of the following treatment it is useful to stick with the general formulation in terms of $\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{P}}(\mathrm{d}x)$ rather than of $M(\tilde{\mathscr{P}})$. We will also assume throughout that (3.12) is satisfied. A key step consists in noting that

$$\mathbb{F}(y; f) := \mathbb{P}\left[\int_{\mathbb{X}} f(x)\tilde{\mathscr{P}}(\mathrm{d}x) \le y\right] = \mathbb{P}\left[\int_{\mathbb{X}} [f(x) - y]\tilde{\mu}(\mathrm{d}x) \le 0\right], \tag{3.13}$$

an idea first employed in [33] for studying Dirichlet means and recalled in (2.22), which clearly applies to NRMIs since it makes use of the representation of the Dirichlet process as normalized gamma process. By virtue of (3.13) the problem of studying a linear functional of a NRMI is reduced to the problem of studying a linear functional of a CRM. And, importantly, the characteristic functions of

linear functionals of CRMs are known and have an elegant Lévy–Khinchine type representation in terms of the underlying Poisson intensity measure. Therefore, by using a suitable inversion formula for characteristic functions one obtains a representation for the probability distribution of $\int_{\mathbb{X}} f(x)\tilde{\mathscr{P}}(\mathrm{d}x)$. In particular, by resorting to Gurland's inversion formula (see Appendix B for details) one obtains

$$\begin{aligned}
&\frac{1}{2}\left\{\mathbb{F}(y;f)+\mathbb{F}(y-0;f)\right\} \qquad (3.14)\\
&=\frac{1}{2}-\frac{1}{\pi}\lim_{\epsilon\downarrow 0,T\uparrow+\infty}\int_{\epsilon}^{T}\frac{1}{t}\operatorname{Im}\mathbb{E}\left[\exp\left\{\mathrm{i}t\int_{\mathbb{X}}[f(x)-y]\tilde{\mu}(\mathrm{d}x)\right\}\right]\mathrm{d}t\\
&=\frac{1}{2}-\frac{1}{\pi}\lim_{\epsilon\downarrow 0,T\uparrow+\infty}\int_{\epsilon}^{T}\frac{1}{t}\operatorname{Im}\exp\left\{-\int_{\mathbb{X}\times\mathbb{R}^+}[1-\mathrm{e}^{\mathrm{i}tv\,(f(x)-y)}]\rho_x(\mathrm{d}v)\alpha(\mathrm{d}x)\right\}\mathrm{d}t
\end{aligned}$$

where $\operatorname{Im} z$ stands for the imaginary part of $z\in\mathbb{C}$. This line of reasoning together with the proof of the absolute integrability in the origin of the integrand in (3.14) led [73] to establish the following result.

Theorem 3.2. *Let $\tilde{\mathscr{P}}$ be a NRMI. Then, for every $y\in\mathbb{R}$ one has*

$$\begin{aligned}
&\frac{1}{2}\left\{\mathbb{F}(y;f)+\mathbb{F}(y-0;f)\right\}\\
&=\frac{1}{2}-\frac{1}{\pi}\lim_{T\uparrow+\infty}\int_{0}^{T}\frac{1}{t}\exp\left\{\int_{\mathbb{X}\times\mathbb{R}^+}[\cos(tv[f(x)-y])-1]\rho_x(\mathrm{d}v)\alpha(\mathrm{d}x)\right\}\\
&\qquad\qquad\times\sin\left(\int_{\mathbb{X}\times\mathbb{R}^+}\sin(tv[f(x)-y])\rho_x(\mathrm{d}v)\alpha(\mathrm{d}x)\right)\mathrm{d}t.
\end{aligned}$$

Now we present three examples, which show how the expression can be simplified for particular choices of NRMIs.

Example 3.4. The first NRMI we consider is the normalized stable process [52], which is obtained by normalizing a stable CRM, i.e. a CRM having Poisson intensity of the form $\rho(\mathrm{d}v)\alpha(\mathrm{d}x)=\sigma v^{-\sigma-1}\,\mathrm{d}v/\Gamma(1-\sigma)\alpha(\mathrm{d}x)$ with $\sigma\in(0,1)$ and α a finite measure. A mean of such a random probability measure is finite if and only if $\int|f(x)|^{\sigma}\,\alpha(\mathrm{d}x)<\infty$, which coincides with the condition (3.4) required for a general $\tilde{\mathscr{P}}_{(\sigma,\theta)}$ mean to be finite. As shown in [73], application of Theorem 3.2 leads to

$$\mathbb{F}(y;f)=\frac{1}{2}-\frac{1}{\pi\sigma}\arctan\left(\frac{\int_{\mathbb{X}}\operatorname{sgn}(f(x)-y)|f(x)-y|^{\sigma}\alpha(\mathrm{d}x)}{\int_{\mathbb{X}}|f(x)-y|^{\sigma}\alpha(\mathrm{d}x)}\tan\frac{\pi\sigma}{2}\right) \qquad (3.15)$$

where sgn stands for the sign function. Such a distribution turns out to be important for the study of the zero–range process introduced in [80]: the zero–range process is a popular model in physics used for describing interacting particles which jump between the sites of a lattice with a rate depending on the occupancy of the departure site. See [22] for a recent review. An interesting case corresponds to the one–dimensional zero–range process acting in a symmetric

medium and the inverse of the particle rates are in the domain of attraction of a strictly σ–stable distribution. Under these assumptions, in [70] it is shown that the associated sequence of *fugacity processes* weakly converges to a process whose marginal distribution, at index point $z \in (0,1)$, is described by (3.15) with $f(x) = \mathbb{1}_{(-\infty,z]}(x)$ and $y - v/(u-v)$ replacing y for $u > v$. It should also be mentioned that Theorem 1 in [70] provides an expression for the moments of $\tilde{\mathscr{P}}((-\infty, z])$ when $\tilde{\mathscr{P}}$ is a normalized σ–stable process. In a similar fashion as for the Dirichlet case examined in Section 2.4, one has an expression involving a sum over a space of partitions of $\{1,\ldots,n\}$ if the n–th moment is to be computed. See also [38] for a general structural expression for moments of means of species sampling models.

Since the σ–stable NRMI is a PD$(\sigma, 0)$ process, the considerations in Section 3.1 establish interesting connections of formula (3.15) with occupation time phenomena for the Brownian motion and more general Bessel processes. In particular, if $f(x) = \mathbb{1}_C(x)$ for some set C in $\mathscr{X}$ such that $\alpha(C)[\alpha(\mathbb{X})]^{-1} = p \in (0,1)$, then the probability distribution of $\tilde{\mathscr{P}}(C)$ coincides with the generalized arcsine laws, originally determined by [54]. It also represents the probability distribution of the time spent positive by a skew Bessel process of dimension $2-2\sigma$ (see, e.g., [68, 69]). It then follows immediately from (3.15) that the cumulative distribution function of Lamperti's generalized arcsine laws can be represented as

$$\mathbb{F}(y; \mathbb{1}_C) = \frac{1}{2} - \frac{1}{\pi\sigma} \arctan\left(\frac{(1-y)^\sigma p - y^\sigma \bar{p}}{(1-y)^\sigma p + y^\sigma \bar{p}} \tan\frac{\pi\sigma}{2} \right). \tag{3.16}$$

for any $y \in (0,1)$. See Example 4.1 in [44] for a different representation of (3.16). By further assuming $\sigma = 1/2$ and some algebra, (3.16) can be further simplified to

$$\mathbb{F}(y; \mathbb{1}_C) = \frac{2}{\pi} \arcsin\left(\frac{\sqrt{y}\bar{p}}{\sqrt{(1-y)p^2 + y\bar{p}^2}} \right).$$

It is then evident that, if also $p = 1/2$, the previous expression coincides with $2\,\pi^{-1} \arcsin(\sqrt{y})$, which is the famous Lévy's arcsine formula for the Brownian motion [55]. □

Example 3.5. An important family of NRMIs, considered in [40, 67, 58], is obtained by setting $\tilde{\mu}$ as a generalized gamma CRM, which is characterized by a Poisson intensity of the type $\rho(\mathrm{d}v)\alpha(\mathrm{d}x) = \sigma[\Gamma(1-\sigma)]^{-1}\, v^{-1-\sigma}\mathrm{e}^{-\tau v}\,\mathrm{d}v\,\alpha(\mathrm{d}x)$ with $\sigma \in (0,1)$, $\tau > 0$ and α a boundedly finite measure on $\mathbb{X}$ (see Example A.2 in Appendix A for details). Given that $\rho(\mathbb{R}^+) = \infty$ and that, for any $\lambda > 0$,

$$\int_{\mathbb{X}\times\mathbb{R}^+} (1 - \mathrm{e}^{-\lambda v})\rho_x(\mathrm{d}v)\alpha(\mathrm{d}x) = (\tau+\lambda)^\sigma \alpha(\mathbb{X}) - \tau^\sigma \alpha(\mathbb{X}),$$

the normalization procedure leads to a well–defined random probability measure if and only if α is a finite measure. The Dirichlet process is recovered by letting $\sigma \to 0$, whereas the normalized σ–stable process [52] arises if $\tau = 0$. Moreover, the resulting NRMI is a normalized inverse Gaussian CRM [56] if $\sigma = 1/2$.

For a generalized gamma NRMI mean, the finiteness condition (3.12) reduces to $\int_{\mathbb{X}}(\tau + |f(x)|)^{\sigma}\alpha(\mathrm{d}x) < \infty$. Note that the Feigin and Tweedie condition (2.4) for the Dirichlet case is recovered by letting $\sigma \to 0$. One can then use Theorem 3.2 for determining also the distribution of such a mean as done in [46]. Recall that $\alpha(\,\cdot\,) = \theta P_0(\,\cdot\,)$ and set $\beta = \theta\tau^{\sigma} > 0$. Then, its expression is given by

$$\mathbb{F}(y; f) = \frac{1}{2} - \frac{\mathrm{e}^{\beta}}{\pi}\int_0^{\infty}\frac{1}{t}\,\mathrm{e}^{-\beta\,A_y(t)}\,\sin(\beta\,B_y(t))\,\mathrm{d}t, \tag{3.17}$$

where

$$A_y(t) = \int_{\mathbb{X}}\left[1 + t^2(f(x)-y)^2\right]^{\frac{\sigma}{2}}\cos\left\{\sigma\arctan\left[t(f(x)-y\right]\right\}P_0(\mathrm{d}x)$$

$$B_y(t) = \int_{\mathbb{X}}\left[1 + t^2(f(x)-y)^2\right]^{\frac{\sigma}{2}}\sin\left\{\sigma\arctan\left[t(f(x)-y)\right]\right\}P_0(\mathrm{d}x)$$

□

Example 3.6. Another interesting NRMI is the normalized extended gamma CRM, which is obtained by normalizing a CRM with intensity $\nu(\mathrm{d}v, \mathrm{d}x) = \mathrm{e}^{-\beta(x)v}\,v^{-1}\,\mathrm{d}v\,\alpha(\mathrm{d}x)$ with β a positive real valued function and α a boundedly finite measure on $\mathbb{X}$ (see Example A.3 for details). Being an infinite activity process, the NRMI is well defined for any combination of β and α satisfying $\int_{\mathbb{X}}\log[1 + (\beta(x))^{-1}]\alpha(\mathrm{d}x) < \infty$. In contrast to the previous two examples, the extended gamma NRMI is non–homogeneous: its random probability masses depend on the locations where they are concentrated. As for means of extended gamma NRMIs, the necessary and sufficient condition for finiteness (3.12) becomes

$$\int_{\mathbb{X}}\log\{1 + |f(x)|\,(\beta(x))^{-1}\}\alpha(\mathrm{d}x) < \infty.$$

One obtains the Feigin and Tweedie condition (2.4) for the Dirichlet case by setting β equal to a constant. As far as the cumulative distribution function of an extended gamma NRMI mean is concerned, note that

$$\mathbb{F}(y; f) = \mathbb{P}\left[\int_{\mathbb{R}}[f(x) - y]\,\tilde{\mu}(\mathrm{d}x) \le 0\right] = \mathbb{P}\left[\int_{\mathbb{R}}\frac{f(x) - y}{\beta(x)}\,\tilde{\gamma}(\mathrm{d}x) \le 0\right]$$

where $\tilde{\mu}$ and $\tilde{\gamma}$ are extended gamma and gamma CRMs, respectively. The latter equality displayed above follows from arguments similar to those in Proposition 3 in [63]. This means that the desired distribution is obtained by inverting the characteristic function of a linear functional of a gamma process, which is exactly the same situation as for the Dirichlet process. Hence, Theorem 3.2 leads to a similar expression to the one given in Theorem 2.4, namely

$$\mathbb{F}(y; f) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty}\frac{1}{t}\,\mathrm{e}^{-\frac{1}{2}\int_{\mathbb{X}}\log\left[1+t^2\left(\frac{f(x)-y}{\beta(x)}\right)^2\right]\alpha(\mathrm{d}x)}\times \\ \times\,\sin\left(\int_{\mathbb{X}}\arctan\left[t\frac{f(x)-y}{\beta(x)}\right]\alpha(\mathrm{d}x)\right)\mathrm{d}t. \tag{3.18}$$

□

Another example of NRMI, where explicit expressions can be obtained, is represented by a generalization of the Dirichlet process obtained by normalizing superposed gamma processes with different scale parameters. See [73] for details.

As for the computation of moments of means of NRMIs, it is possible to obtain a generalization of the formula displayed in (2.30). In order to do so we introduce the following quantity

$$\Pi_k^{(n)}(n_1,\ldots,n_k;f)=\frac{1}{\Gamma(n)}\int_0^\infty u^{n-1}\,\mathrm{e}^{-\int_{\mathbb{X}\times\mathbb{R}^+}(1-\mathrm{e}^{-us})\rho_x(\mathrm{d}s)\,\alpha(\mathrm{d}x)}\times \\ \times\prod_{i=1}^k\left\{\int_{\mathbb{X}}\alpha(\mathrm{d}x)\,f^{n_i}(x)\int_0^\infty s^{n_i}\mathrm{e}^{-us}\rho_x(\mathrm{d}s)\right\}\mathrm{d}u \quad (3.19)$$

which is defined one the set of positive integers $\{\boldsymbol{n}=(n_1,\ldots,n_k): n_i\geq 1, |\boldsymbol{n}|=n\}$, where $|\boldsymbol{n}|=\sum_{i=1}^k n_i$. Note that when $f\equiv 1$, then $\Pi_k^{(n)}(n_1,\ldots,n_k;f)=\Pi_k^{(n)}(n_1,\ldots,n_k)$ describes the so-called *exchangeable partition probability function* induced by a NRMI and derived in [45]. Hence, one has the following statement

Theorem 3.3. *Let $\tilde{\mathscr{P}}$ be a NRMI. If $f:\mathbb{X}\to\mathbb{R}$ is a measurable function for which (3.12) holds true and if the n–th moment of $\int f\,\mathrm{d}\tilde{\mathscr{P}}$ is finite, then*

$$\mathbb{E}\left[\left(\int_{\mathbb{X}}f(x)\,\tilde{\mathscr{P}}(\mathrm{d}x)\right)^n\right]=\sum_{k=1}^n\sum_{(m_1,\ldots,m_n)\in Z(n,k)}\frac{n!}{\prod_{j=1}^n(j!)^{m_j}m_j!}\times \\ \times\Pi_k^{(n)}(\underbrace{1,\ldots,1}_{m_1},\underbrace{2,\ldots,2}_{m_2},\ldots,\underbrace{n,\ldots,n}_{m_n};f), \quad (3.20)$$

where the set $Z(n,k)$ is defined in (2.12).

Note that when the NRMI $\tilde{\mathscr{P}}$ is generated by normalizing a homogeneous CRM, i.e. $\rho_x=\rho$ for any x in $\mathbb{X}$, then the formula above reduces to

$$\mathbb{E}\left[\left(\int_{\mathbb{X}}f(x)\,\tilde{\mathscr{P}}(\mathrm{d}x)\right)^n\right]=\sum_{k=1}^n\sum_{(m_1,\ldots,m_n)\in Z(n,k)}\frac{n!}{\prod_{j=1}^n(j!)^{m_j}m_j!}\times \\ \times\Pi_k^{(n)}(\underbrace{1,\ldots,1}_{m_1},\underbrace{2,\ldots,2}_{m_2},\ldots,\underbrace{n,\ldots,n}_{m_n})\prod_{i=1}^n\left(\int_{\mathbb{X}}f^i(x)\alpha(\mathrm{d}x)\right)^{m_i}. \quad (3.21)$$

Since homogeneous NRMI are a subclass of special sampling models [66], (3.21) is an instance of where the structural representation for the moments of species sampling means, derived in [38], can be made explicit. The result above is in the spirit of Lo's treatment of mixtures of the Dirichlet process in [60]. See also [40]. The following examples show how (3.20) can be easily applied, even if the

concrete evaluation of the n–th moment is a challenging task since one has to deal with sums over spaces of partitions.

Example 3.7. Since the Dirichlet process is a special case of homogeneous NRMI, one can recover (2.30) from (3.21). Suppose $\alpha = \theta P_0$ and note that in this case $\rho_x(\mathrm{d}s) = \rho(\mathrm{d}s) = \mathrm{e}^{-s}s^{-1}\,\mathrm{d}s$. Hence

$$\Pi_k^{(n)}(n_1,\ldots,n_k) = \frac{\theta^k}{(\theta)_n}\prod_{i=1}^{k}(n_i-1)!$$

will yield to the representation in terms of partial exponential Bell polynomials as recalled in Theorem 2.5. □

Example 3.8. Suppose $\tilde{\mathscr{P}}$ is a stable NRMI already considered in Example 3.4. Without loss of generality, let α be a probability measure on $\mathbb{X}$. In this case

$$\Pi_k^{(n)}(n_1,\ldots,n_k) = \frac{\sigma^{k-1}\Gamma(k)}{\Gamma(n)}\prod_{i=1}^{k}(1-\sigma)_{n_i-1}$$

so that the n–th moment of $\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{P}}(\mathrm{d}x)$ coincides with

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{\mathbb{X}} f(x)\,\tilde{\mathscr{P}}(\mathrm{d}x)\right)^n\right]\\
&\quad= (n-1)\sum_{k=1}^{n}\Gamma(k)\,\sigma^{k-1}\sum_{(m_1,\ldots,m_n)\in Z(n,k)}\frac{1}{\prod_{j=1}^{n}(j!)^{m_j}\,m_j!}\prod_{i=1}^{n}((1-\sigma)_{i-1}\,r_{i,\alpha})^{m_i}\\
&\quad= \sum_{k=1}^{n}\frac{\Gamma(k)}{\Gamma(n)}\,\sigma^{k-1}\,B_{n,k}(r_{1,\alpha},(1-\sigma)_{2-1}\,r_{2,\alpha},\ldots,(1-\sigma)_{n-k}\,r_{n-k+1,\alpha}).
\end{aligned}$$

If in the previous expression one sets $f = \mathbb{1}_{(-\infty,x]}$ and α is a probability measure on $(0,1)$, then one obtains the formula described in Theorem 1 of [70]. □

Having determined the probability distribution of $\int f\,\mathrm{d}\tilde{\mathscr{P}}$, a Bayesian is interested in the evaluation of the posterior distribution, given a sample of n observations. In the Dirichlet case, the determination of the prior distribution of $M(\tilde{\mathscr{D}}_\alpha)$ is enough to solve the issue because of conjugacy $\tilde{\mathscr{D}}_\alpha$. Hence the posterior distribution of $M(\tilde{\mathscr{D}}_\alpha)$, given $X_1,\ldots,X_n$, coincides with the distribution of $M(\tilde{\mathscr{D}}_{\alpha^*})$, with $\alpha^* = \alpha + \sum_{i=1}^{n}\delta_{X_i}$. If the prior is not conjugate, further efforts are needed for determining the posterior distribution of a linear functional. This is the case of NRMIs different from the Dirichlet process: indeed, as shown in [43] they are not conjugate. Nonetheless, for NRMIs a full description of the posterior distribution can be provided. Here, we briefly summarize the results achieved in [73, 46].

Assume $(X_n)_{n\geq 1}$ to be a sequence of exchangeable observations, defined on $(\Omega,\mathscr{F},\mathbb{P})$ and with values in $\mathbb{X}$, such that, given a NRMI $\tilde{\mathscr{P}}$, the X_i's are i.i.d. with distribution $\tilde{\mathscr{P}}$, i.e. for any $B_i\in\mathscr{X}$, $i=1,\ldots,n$ and $n\geq 1$

$$\mathbb{P}\left[X_1\in B_1,\ldots,X_n\in B_n\,\middle|\,\tilde{\mathscr{P}}\right] = \prod_{i=1}^{n}\tilde{\mathscr{P}}(B_i).$$

Since a NRMI selects a.s. discrete distributions, ties in the observations will appear with positive probability. Set $\boldsymbol{X} = (X_1, \ldots, X_n)$ and clearly one can always represent $\boldsymbol{X}$ as $(\boldsymbol{X}^*, \boldsymbol{\pi})$, where $\boldsymbol{X}^* = (X_1^*, \ldots, X_{n(\boldsymbol{\pi})}^*)$ denotes the distinct observations within the sample and $\boldsymbol{\pi} = \{C_1, \ldots, C_{n(\boldsymbol{\pi})}\}$ stands for the corresponding partition of the integers $\{1, \ldots, n\}$ recording which observations within the sample are equal, that is $C_j = \{i : X_i = X_j^*\}$. The number of elements in the j-th set of the partition is indicated by n_j, for $j = 1, \ldots, n(\boldsymbol{\pi})$, so that $\sum_{j=1}^{n(\boldsymbol{\pi})} n_j = n$. For the remainder of the section we will deal with NRMIs derived from CRMs with non–atomic base measure α in (A.5).

Before dealing with functionals of a NRMIs, it is useful to first recall their posterior characterization given in [45]. For any pair of random elements Z and W defined on $(\Omega, \mathscr{F}, \mathbb{P})$, we use the symbol $Z^{(W)}$ to denote a random element on $(\Omega, \mathscr{F}, \mathbb{P})$ whose distribution coincides with a regular conditional distribution of Z, given W. Moreover, denote the Laplace exponent of a CRM by $\psi(f) := \int_{\mathbb{R}^+ \times \mathbb{X}} [1 - \mathrm{e}^{-v\, f(x)}] \rho_x(\mathrm{d}v) \alpha(\mathrm{d}x)$ for any measurable function $f : \mathbb{X} \to \mathbb{R}$ for which the integrability condition (3.12) holds true. Introduce now a latent variable, denoted by U_n, whose conditional distribution, given $\boldsymbol{X}$, admits a density function (with respect to the Lebesgue measure on $\mathbb{R}$) coinciding with

$$f_{U_n}^{\boldsymbol{X}}(u) \propto u^{n-1} \prod_{i=1}^{n(\boldsymbol{\pi})} \tau_{n_i}(u | X_i^*)\, \mathrm{e}^{-\psi(u\mathbf{1})} \tag{3.22}$$

where $\mathbf{1}(x) = 1$, for any $x \in \mathbb{X}$, and

$$\tau_{n_i}(u | X_i^*) = \int_{\mathbb{R}^+} s^{n_i}\, \mathrm{e}^{-us}\, \rho_{X_i^*}(\mathrm{d}s) \tag{3.23}$$

for $i = 1, \ldots, n(\boldsymbol{\pi})$. Indeed, the posterior distribution, given $\boldsymbol{X}$, of the CRM $\tilde{\mu}$ defining a NRMI is a mixture with respect to the distribution of the latent variable U_n. Specifically,

$$\tilde{\mu}^{(U_n, \boldsymbol{X})} \stackrel{d}{=} \tilde{\mu}^{(U_n)} + \sum_{i=1}^{n(\boldsymbol{\pi})} J_i^{(U_n, \boldsymbol{X})}\, \delta_{X_i^*}$$

where: $\tilde{\mu}^{(U_n)}$ is a CRM with intensity

$$\nu^{(U_n)}(\mathrm{d}s, \mathrm{d}x) = \mathrm{e}^{-U_n s}\, \rho_x(\mathrm{d}s)\, \alpha(\mathrm{d}x); \tag{3.24}$$

the X_i^*'s are the fixed points of discontinuity; the $J_i^{(U_n, \boldsymbol{X})}$'s are the corresponding jumps, which are mutually independent and independent from $\tilde{\mu}^{(U_n)}$, and whose density is given by

$$f_{J_i}^{(U_n, \boldsymbol{X})}(s) \propto s^{n_i} \mathrm{e}^{-U_n s} \rho_{X_i^*}(\mathrm{d}s). \tag{3.25}$$

See [45, Theorem 1] for details. Finally, denote by $\psi^{(u)}$ and $J_r^{(u, \boldsymbol{X})}$ the Laplace exponent of the CRM defined by (3.24) and the jumps whose density is given by (3.25), respectively, with $U_n = u$.

Now we are in a position to provide a description of the exact posterior distribution of a linear functional of a general NRMI. First of all, it is to be noted that such a distribution is always absolutely continuous w.r.t. the Lebesgue measure $\mathbb{R}$.

Theorem 3.4. *Let $\tilde{\mathscr{P}}$ be a NRMI. Then, the posterior density function of $\int_{\mathbb{X}} f(x)\tilde{\mathscr{P}}(\mathrm{d}x)$, given $\boldsymbol{X}$, is of the form*

$$\varrho^{\boldsymbol{X}}(y;f)=\begin{cases}\displaystyle\int_0^{\infty}\mathrm{Re}\{\chi_f(t,y)\}\,\mathrm{d}t & \text{if } n=1\\ \displaystyle(-1)^{p+1}\int_{-\infty}^{y}\int_0^{\infty}[(y-z)\,t]^{n-1}\mathrm{Im}\{\chi_f(t,z)\}\,\mathrm{d}t\,\mathrm{d}z & \text{if } n=2p\\ \displaystyle(-1)^{p}\int_{-\infty}^{y}\int_0^{\infty}[(y-z)\,t]^{n-1}\mathrm{Re}\{\chi_f(t,z)\}\,\mathrm{d}t\,\mathrm{d}z & \text{if } n=2p+1\end{cases}$$

where $p\geq 1$

$$\chi_f(t,z)=\frac{\mathrm{e}^{-\psi(-\mathrm{i}t(f-z\mathbf{1}))}\prod_{j=1}^{n(\boldsymbol{\pi})}\kappa_{n_j}\left(\mathrm{i}t\left[f(X_j^*)-z\right]\,\middle|\,X_j^*\right)}{\pi\int_0^{+\infty}u^{n-1}\left[\prod_{j=1}^{n(\boldsymbol{\pi})}\tau_{n_j}(u|X_j^*)\right]\mathrm{e}^{-\psi(u\mathbf{1})}\,\mathrm{d}u},$$

$\kappa_{n_j}(\mathrm{i}t[f(X_j^*)-z]|X_j^*)=\int_0^{+\infty}v^{n_j}\,\mathrm{e}^{\mathrm{i}tv(f(X_j^*)-z)}\rho_{X_j^*}(\mathrm{d}v)$ *and* $\tau_{n_j}(u|X_j^*)$ *is as in* (3.23)*, for $j=1,\ldots,n(\boldsymbol{\pi})$. Moreover, the posterior cumulative distribution function of $\int_{\mathbb{X}} f(x)\tilde{\mathscr{P}}(\mathrm{d}x)$, given $\boldsymbol{X}$, can be expressed as*

$$\mathbb{F}^{\boldsymbol{X}}(y;\,f)=\frac{1}{2}-\frac{1}{\pi}\lim_{T\to+\infty}\int_0^T\frac{1}{t}\int_0^{+\infty}\zeta_f(y;\,u,t)\,f_{U_n}^{\boldsymbol{X}}(u)\,\mathrm{d}u\,\mathrm{d}t$$

where

$$\zeta_f(y;\,u,t):=\mathrm{Im}\left\{\mathrm{e}^{-\psi^{(u)}(-\mathrm{i}t(f-y\mathbf{1}))}\,\mathbb{E}\left[\mathrm{e}^{\mathrm{i}t\sum_{r=1}^{n(\boldsymbol{\pi})}(f(X_r^*)-y)J_r^{(u,\boldsymbol{X})}}\right]\right\},$$

and $f_{U_n}^{\boldsymbol{X}}$ is the density of the latent variable U_n given in (3.22).

The proof of this result heavily relies on the useful and powerful techniques and tools developed in [74]. As for the determination of the posterior density, the starting point is given by its exact representation in the case of NRMIs with α in (A.5) having finite support [73], which, loosely speaking, handles observations as derivatives of the characteristic function. In order to achieve the representation presented above, one needs to exploit the fact that α is non–atomic in connection with a martingale convergence argument. The cumulative distribution function, instead, arises by combining Theorem 3.2 with the posterior representation of NRMIs.

The reader is referred to [73, 46, 57] for expressions of the posterior distribution of functionals based on particular NRMIs. Here, we just point out that, by the application of the previous results, in [73] a particularly simple formula for

the density of Dirichlet means has been obtained. Indeed, provided there exists an x such that $\alpha(\{x\}) \geq 1$, one has

$$\varrho(y,f) = \frac{\alpha(\mathbb{X})-1}{\pi}\int_0^{+\infty} \mathrm{Re}(\mathrm{e}^{-\int_{\mathbb{X}} \log(1-\mathrm{i}t(f(x)-y))\alpha(\mathrm{d}x)})\mathrm{d}t.$$

We close this section by considering linear functionals of hierarchical mixture models, which is nowadays the most common use of Bayesian nonparametric procedures. Letting $\mathbb{Y}$ be a Polish space equipped with the Borel σ-algebra $\mathscr{Y}$, one defines a random density (absolutely continuous with respect to some σ-finite measure λ on $\mathbb{Y}$) driven by a random discrete distribution i.e.

$$\tilde{p}(y) = \int_{\mathbb{X}} k(y,x)\,\tilde{\mathscr{P}}(\mathrm{d}x) \tag{3.26}$$

where k is a density function on $\mathbb{Y}$ indexed by some parameter with values in $\mathbb{X}$. A typical choice for k is represented by the density function of normal distribution: in such a case $\tilde{\mathscr{P}}$ controls the means (and possibly also the variances) of the random mixture density. This approach is due to Lo [60] who defined a random density as in (3.26) with $\tilde{\mathscr{P}}$ being the Dirichlet process: this model is now commonly referred to as mixture of Dirichlet process (MDP). One can obviously replace the Dirichlet process with any NRMI: interesting behaviours especially in terms of the induced clustering mechanism appear for particular choices of NRMIs. See, e.g., [56, 58, 57]. The study of mean functionals of such models has been addressed in [63, 46]. The first thing to note is that, as far as the prior distribution is concerned, the problem of studying a linear functional of a NRMI mixture can be reduced to studying (different) linear functional of a simple NRMI. Indeed, for a NRMI mixture density $\tilde{p}$ with $\tilde{\mathscr{P}}$ being the corresponding NRMI, one has

$$\int_{\mathbb{Y}} f(y)\,\tilde{p}(y)\lambda(\mathrm{d}y) = \int_{\mathbb{X}} g(x)\,\tilde{\mathscr{P}}(\mathrm{d}x) \tag{3.27}$$

where $g(x) = \int_{\mathbb{Y}} f(y)k(y,x)\lambda(\mathrm{d}y)$. Hence, as shown in [63], a necessary and sufficient condition for $\int_{\mathbb{Y}} f(y)\,\tilde{p}(y)\lambda(\mathrm{d}y)$ being a.s. finite becomes

$$\int_{\mathbb{X}\times\mathbb{R}^+} \left(1-\mathrm{e}^{-tvg^*(x)}\right)\,\rho_x(\mathrm{d}v)\,\alpha(\mathrm{d}x) < +\infty \qquad \forall t>0, \tag{3.28}$$

with $g^*(x) = \int_{\mathbb{Y}} |f(y)|k(y,x)\lambda(\mathrm{d}y)$. For instance, in the MDP case (3.28) reduces to

$$\int_{\mathbb{X}} \log[1+g^*(x)]\,\alpha(\mathrm{d}x) < +\infty. \tag{3.29}$$

From (3.27), it follows the prior distribution of the mean of a mixture can be evaluated by applying Theorem 3.2. As far as the posterior distribution of NRMI mixtures is concerned, general expressions for their density and cumulative distribution function can be found in [46]. For the popular MDP case, which is also not conjugate, an expression for the posterior density function has been first obtained in [63].

Theorem 3.5. *Let $\tilde{p}$ be a random density defined as in (3.26) with $\tilde{\mathscr{P}} = \tilde{\mathscr{D}}_\alpha$ and f be a measurable function satisfying (3.29). Then the posterior density function of a mean $\int f(y)\tilde{p}(y)\lambda(\mathrm{d}y)$, given $\boldsymbol{Y}$, is of the form*

$$\varrho^{\boldsymbol{Y}}(y,f) = \frac{\sum_{\boldsymbol{\pi}} \prod_{j=1}^{n(\boldsymbol{\pi})} (n_j - 1)! \int_{\mathbb{X}^{n(\boldsymbol{\pi})}} \varrho^n(y,f) \prod_{i \in C_j} k(Y_i, x_j) \alpha(\mathrm{d}x_j)}{\sum_{\boldsymbol{\pi}} \prod_{j=1}^{n(\boldsymbol{\pi})} (n_j - 1)! \int_{\mathbb{X}} \prod_{i \in C_j} k(Y_i, x) \alpha(\mathrm{d}x)},$$

where

$$\varrho^n(y,f) = \frac{\alpha(\mathbb{X}) + n - 1}{\pi} \int_{\mathbb{R}^+} \mathrm{Re}\{\mathrm{e}^{-\int_{\mathbb{X}} \log(1 - \mathrm{i}t(g(x)-y)) \alpha_n^{\boldsymbol{X}}(\mathrm{d}s)}\} \mathrm{d}t$$

with $\alpha_n^{\boldsymbol{X}}(\,\cdot\,) = \alpha(\,\cdot\,) + \sum_{j=1}^{n(\boldsymbol{\pi})} n_j \delta_{x_j}$ and $g(x) = \int_{\mathbb{Y}} f(y) k(y,x) \lambda(\mathrm{d}y)$.

See [63, 46] for further details and various examples.

Finally, an alternative use of discrete nonparametric priors, such as NRMI, for inference with continuous data is represented by histogram smoothing. Such a problem can be handled by exploiting the so–called "filtered-variate" random probability measures as defined by [16], which essentially coincide with suitable means of random probability measures.

3.3. Means of neutral to the right priors

A first remarkable generalization of the Dirichlet process that has appeared in the literature is due to Doksum [17]. A random distribution function $\tilde{F} = \{\tilde{F}(t) : t \geq 0\}$ on $\mathbb{R}^+$ is said *neutral to the right* (NTR) if, for any choice of points $0 < t_1 < t_2 < \cdots < t_n < \infty$ and for any $n \geq 1$, the random variables

$$\tilde{F}(t_1), \frac{\tilde{F}(t_2) - \tilde{F}(t_1)}{1 - \tilde{F}(t_1)}, \ldots, \frac{\tilde{F}(t_n) - \tilde{F}(t_{n-1})}{1 - \tilde{F}(t_{n-1})}$$

are independent. An important characterization is given in Theorem 3.1 of [17] where it is shown that $\tilde{F}$ is NTR if and only if there exists a CRM $\tilde{\mu}$ on $\mathbb{R}^+$ with $\lim_{t \to \infty} \tilde{\mu}((0,t]) = \infty$ almost surely such that $\tilde{F} = \{\tilde{F}(t) : t \geq 0\}$ coincides in distribution with $\{1 - \mathrm{e}^{-\tilde{\mu}((0,t])} : t \geq 0\}$.

Example 3.9. As recalled in Section 2, the Dirichlet process can also be seen as a neutral to the right process. Indeed, let $\tilde{\mu}$ be a CRM with Laplace transform

$$\begin{aligned} \mathbb{E}\left[\mathrm{e}^{-\lambda \tilde{\mu}((0,t])}\right] &= \exp\left\{-\int_0^t \int_0^\infty \left[1 - \mathrm{e}^{-\lambda s}\right] \frac{\mathrm{e}^{-s\,\alpha((x,\infty))}}{1 - \mathrm{e}^{-s}} \mathrm{d}s\, \alpha(\mathrm{d}x)\right\} \qquad (3.30) \\ &= \exp\left\{-\int_0^\infty \left[1 - \mathrm{e}^{-\lambda s}\right] \frac{(1 - \mathrm{e}^{-s\,\alpha((0,t])})\, \mathrm{e}^{-\alpha((x,\infty))s}}{(1 - \mathrm{e}^{-s})\, s} \mathrm{d}s\right\} \end{aligned}$$

for some finite measure α and for any $\lambda > 0$. The random probability measure whose distribution function is defined by $\tilde{F}(t) = 1 - \mathrm{e}^{-\tilde{\mu}((0,t])}$ coincides with $\tilde{\mathscr{D}}_\alpha$. See [27]. □

Since $\int_{\mathbb{R}^+} x\,\mathrm{d}\tilde{F}(x) = \int_0^\infty \left[1-\tilde{F}(t)\right]\,\mathrm{d}t$, one has that, if $\tilde{F}$ is NTR, then

$$\int_{\mathbb{R}^+} x\,\mathrm{d}\tilde{F}(x) \stackrel{d}{=} \int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t. \tag{3.31}$$

The random quantity appearing on the right–hand–side of (3.31) is known as *exponential functional* and it has attracted great interest in probability and finance. Indeed, it appears in risk theory and in models for the pricing of Asian options and for the determination of the law of a perpetuity. In mathematical physics it represents a key quantity for the analysis of one–dimensional diffusion in a random Lévy environment. See [2] for a comprehensive review and relevant references.

The analysis of distributional properties of $\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t$ is a demanding task. There are still many open problems in this area that certainly deserve a considerable amount of work. Here we confine ourselves to recalling some known results concerning the existence of means of neutral to the right processes and some interesting distributional characterizations. Moreover, we confine ourselves to considering the case where $\tilde{\mu}$ does not have fixed points of discontinuity.

As far as the first issue is concerned, according to Theorem 1 of [2] the condition $\mathbb{P}[\lim_{t\to\infty}\tilde{\mu}((0,t]) = \infty] = 1$ is equivalent to almost sure finiteness of $\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t$ when $\tilde{\mu}$ is a CRM with intensity measure $\nu(\mathrm{d}s,\mathrm{d}x) = \mathrm{d}x\,\rho(\mathrm{d}s)$. We are not aware of necessary and sufficient condition for the existence of $\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t$ for a general CRM $\tilde{\mu}$. In [20], for the case of a homogeneous $\tilde{\mu}$, a sufficient condition for its existence is given in terms of the Lévy intensity. For a non–homogeneous CRM, finiteness of the mean can be guaranteed by assuming the finiteness of its first moment.

An early contribution to the determination of moments of any order of $\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t$ is contained in [9] and it is related to a random Dirichlet distribution function $\tilde{F}$. Indeed, observe that

$$\left(\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t\right)^n = \left\{\int_0^\infty [1-\tilde{F}(s)]\,\mathrm{d}s\right\}^n = \\ = n!\int_0^\infty \mathrm{d}x_1 \int_{x_1}^\infty \mathrm{d}x_2 \cdots \int_{x_{n-1}}^\infty \mathrm{d}x_n \prod_{j=1}^n \left[1-\tilde{F}(x_j)\right]$$

and $\tilde{F}(x_1) \le \cdots \le \tilde{F}(x_n)$. Consequently, the random vector $(\tilde{F}(x_1), \tilde{F}(x_2) - \tilde{F}(x_1), \ldots, \tilde{F}(x_n) - \tilde{F}(x_{n-1}))$ has an n–variate Dirichlet distribution with parameters $(\alpha_1,\ldots,\alpha_{n+1})$, with $\alpha_j = \alpha((x_{j-1},x_j])$ for any $j = 1,\ldots,n$ and $\alpha_{n+1} = \alpha((x_n,\infty))$, and

$$\mathbb{E}\left[\prod_{j=1}^n \left(1-\tilde{F}(x_j)\right)\right] = \frac{\prod_{j=1}^n \{\alpha((x_j\infty)) + n - j\}}{(\theta)_n}$$

where $\theta = \alpha_1 + \cdots + \alpha_{n+1}$. The same line of reasoning can be adapted to a general neutral to the right prior, even if the finite–dimensional distributions of

$\tilde{F}$ are not known. Suppose $\tilde{\mu}$ is a CRM with intensity $\nu(\mathrm{d}s, \mathrm{d}x) = \alpha(\mathrm{d}x)\,\rho_x(\mathrm{d}s)$ and let $\tilde{F}(t) = 1 - \mathrm{e}^{-\tilde{\mu}((0,t])}$. As shown in [20], in this more general case a sufficient condition for the existence of the n–th moment of the mean is

$$\int_0^\infty \exp\left\{-\int_0^{t^{1/n}} \alpha(\mathrm{d}x) \int_0^\infty [1-\mathrm{e}^{-s}]\,\rho_x(\mathrm{d}s)\right\} \mathrm{d}t < \infty$$

and

$$\mathbb{E}\left[\left(\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t\right)^n\right] = n! \int_0^\infty \cdots \int_{x_{n-1}}^\infty \mathrm{e}^{-\sum_{j=1}^n \Phi(x_j)}\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \quad (3.32)$$

where $\Phi(x_j) = \int_{(0,x_j)} \alpha(\mathrm{d}x) \int_0^\infty [1-\mathrm{e}^{-s}]\,\mathrm{e}^{-(n-j)s}\,\rho_x(\mathrm{d}s)$. Note that if $\rho_x = \rho$ and $\alpha(\mathrm{d}x) = \mathrm{d}x$, then

$$\mathbb{E}\left[\left(\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\,\mathrm{d}t\right)^n\right] = \frac{n!}{\prod_{j=1}^n \int_0^\infty [1-\mathrm{e}^{-js}]\,\rho(\mathrm{d}s)}$$

which coincides with a result proved in [6].

Example 3.10. A popular neutral to the right process is given by the beta-Stacy process introduced in [86], which, from a Bayesian point of view, has the appealing feature of being conjugate with respect to both exact and right–censored observations. The CRM corresponding to such a process via the exponential transformation is the log–beta CRM, whose Poisson intensity measure is of the form

$$\nu(\mathrm{d}s, \mathrm{d}x) = \frac{1}{1-e^{-s}}\,\mathrm{e}^{-s\beta(x)}\,\mathrm{d}s\,\alpha(\mathrm{d}x)$$

where β is a positive function, α is a measure concentrated on $\mathbb{R}^+$ which is absolutely continuous with respect to the Lebesgue measure and $\int_0^\infty (\beta(x))^{-1}\,\alpha(dx) = +\infty$. The assumption of absolute continuity of α is equivalent to assuming no fixed points of discontinuity. Clearly, if α is a finite measure and $\beta(x) = \alpha((x, +\infty))$, we recover the process (3.30), which corresponds to the Dirichlet process.

For the mean of a beta–Stacy process the condition for the existence of the moment of order n reduces to

$$\int_{(0,+\infty)} \exp\left\{-\int_{(0,t^{1/n})} \{\beta(s)\}^{-1}\,\alpha(ds)\right\} dt < +\infty$$

and the moment of order n of the mean is

$$n! \int_0^{+\infty} \cdots \int_{t_{n-1}}^{+\infty} \prod_{j=1}^n \exp\left(-\int_{(0,t_j)} \frac{\alpha(ds)}{\beta(s)+n-j}\right) dt_n \cdots dt_1.$$

Given the conjugacy property of the beta-Stacy process, one can also obtain moments of the posterior mean. In fact, let $Y_1, \ldots, Y_N$ denote exchangeable

observations and $S_1, \ldots, S_r$ be the $r \leq N$ exact observations among the Y_j's, the others being right–censored. For simplicity, assume that the observations are all distinct. The posterior distribution of the beta–Stacy process is still a beta–Stacy process. In particular, the corresponding CRM is a log–beta process with fixed points of discontinuities, whose Laplace transform is

$$E(\mathrm{e}^{-\lambda\,\tilde{\mu}((0,t])}) = \exp\left\{-\int_0^t \int_0^\infty (1-\mathrm{e}^{-\lambda s})\nu^*(\mathrm{d}s,\mathrm{d}x)\right\} \prod_{\{k:\, S_k \leq t\}} E\left(\mathrm{e}^{-\lambda J_k}\right), \tag{3.33}$$

with

$$\nu^*(\mathrm{d}s,\mathrm{d}x) = \frac{1}{1-\mathrm{e}^{-s}}\, e^{-s\{\beta(x)+M(x)\}}\,\mathrm{d}s\,\alpha(\mathrm{d}x),$$

where $M(\,\cdot\,) := \sum_{i=1}^N \mathbb{1}_{[\,\cdot\,,+\infty)}(Y_i)$. Moreover, $S_1, \ldots, S_r$ are the fixed points of discontinuity corresponding to the exact observations, the random jump J_i being exponentially distributed with parameter $\tau_i := \beta(S_i) + M(S_i) - 1$ for $i = 1, \ldots, r$. See [86] for details.

Hence, as an expression for the posterior moments of the mean of a beta–Stacy process one obtains

$$n! \int_0^{+\infty} \int_{t_1}^{+\infty} \cdots \int_{t_{n-1}}^{+\infty} \prod_{j=1}^n \exp\left(-\int_{(0,t_j)} \frac{\alpha(dx)}{\beta(x)+M(x)+n-j}\right) \times \tag{3.34}$$
$$\times \left(\prod_{\{i:\, S_i \leq t_j\}} \frac{\tau_i + n - j}{\tau_i + n - j + 1}\right) dt_n \cdots dt_1$$

with the proviso that, if $\{i :\, S_i \leq t_j\} = \emptyset$ for some j, then $\prod_\emptyset := 1$. □

Unlike other means of random probability measures discussed so far, the problem of determining the distribution of a mean of a neutral to the right process has not found a solution with the exception of the Dirichlet process; however, it is to be noted that in order to achieve these results representations of the Dirichlet process different from the one recalled in Example 3.9 were used. The only available characterization has been given in [6] and it can be summarized in the following theorem.

Theorem 3.6. *Suppose $\tilde{\mu}$ is a CRM with Lévy intensity $\nu(\mathrm{d}s,\mathrm{d}x) = \mathrm{d}x\,\rho(\mathrm{d}s)$ such that $\int_0^\infty \rho((s,\infty))\,\mathrm{d}s < \infty$. Then the probability distribution of $\int_0^\infty \mathrm{e}^{-\tilde{\mu}((0,t])}\mathrm{d}t$ is absolutely continuous with respect to the Lebesgue measure with density q solving the integral equation*

$$q(y) = \int_y^\infty \mathrm{d}z \int_{\log(z/y)}^\infty \rho(\mathrm{d}s)\; q(z). \tag{3.35}$$

Unfortunately, so far a solution for the integral equation (3.35) is not known and so the determination of the density q still remains an open problem.

4. Concluding remarks

In this paper we have tried to give a broad account on distributional properties of means of random probability measures and to relate contributions from different areas. There are still several interesting open issues about means and among them we mention: the study of relation between the distribution of the random mean and the expected value of the corresponding random probability measure for processes different from the Dirichlet process such as, e.g., the normalized stable process; the determination of the posterior distribution of means of processes different from NRMI and, most notably, of the two parameter Poisson–Dirichlet process; the derivation of closed form expressions for the density of means of neutral to the right processes, at least in some specific cases; the derivation of analogues of the Cifarelli–Regazzini identity for processes different than the two parameter Poisson–Dirichlet process; further investigations of the interplay between means and excursion theory, combinatorics and special functions.

The obvious and important following step consists in studying non–linear functionals of random probability measures. As for quadratic functionals, various preliminary results are already available for the variance functional of a Dirichlet process $V(\tilde{\mathscr{D}}_\alpha) = \int x^2 \tilde{\mathscr{D}}_\alpha(\mathrm{d}x) - [\int x \tilde{\mathscr{D}}_\alpha(\mathrm{d}x)]^2$: the moments and Laplace transform of $V(\tilde{\mathscr{D}}_\alpha)$ have been obtained in [7], expressions for its cumulative distribution function and density function are given in [72], an alternative representation of its Laplace transform has been derived in [59], whereas in [19] a stochastic equation for $V(\tilde{\mathscr{D}}_\alpha)$ is derived. It is to be mentioned that, in the above papers, the variance functional is treated as a linear combination of the coordinates of the vector of means $(\int x^2 \tilde{\mathscr{P}}(\mathrm{d}x), \int x \tilde{\mathscr{P}}(\mathrm{d}x))$: hence, the determination of distributional properties of genuinely quadratic functionals such as, e.g., the mean difference $\int |x-y| \tilde{\mathscr{P}}(\mathrm{d}x)\tilde{\mathscr{P}}(\mathrm{d}y)$, is still an open problem. Results for non–linear functionals such as the quantiles are provided, for the Dirichlet case, in [35, 37]. Finally, representations for square integrable functionals of the Dirichlet process in terms of an infinite orthogonal sum of multiple integrals of increasing order are derived in [64].

Appendix A: Completely random measures

In this Appendix we provide a concise account on completely random measures, a concept introduced by Kingman [51], which allows to unify in an elegant way the classes of random probability measures dealt with in the previous sections: indeed, all of them can be derived as suitable transformations of completely random measures.

Let $(\mathbb{X}, \mathscr{X})$ be a Polish space equipped with the corresponding Borel σ–field and recall that a measure μ on $\mathbb{X}$ is said to be boundedly finite if $\mu(A) < +\infty$ for every bounded measurable set A. Denote by $(\mathcal{M}_\mathbb{X}, \mathscr{M}_\mathbb{X})$ the space of boundedly finite measures endowed with the corresponding Borel σ–algebra. See [14] for an exhaustive account. Let now $\tilde{\mu}$ be a measurable mapping from $(\Omega, \mathscr{F}, \mathbb{P})$ into

$(\mathcal{M}_{\mathbb{X}}, \mathscr{M}_{\mathbb{X}})$ and such that for any $A_1, \ldots, A_n$ in $\mathscr{X}$, with $A_i \cap A_j = \varnothing$ for $i \neq j$, the random variables $\tilde{\mu}(A_1), \ldots, \tilde{\mu}(A_n)$ are mutually independent. Then $\tilde{\mu}$ is termed *completely random measure* (CRM).

The reader is referred to [53] for a detailed treatment of the subject. Here, we confine ourselves to highlighting some aspects of CRMs, which are essential for the study of mean functionals. A CRM on $\mathbb{X}$ can always be represented as the sum of two components: a proper CRM $\tilde{\mu}_c = \sum_{i=1}^{\infty} J_i \delta_{Y_i}$, where both the positive jumps J_i's and the $\mathbb{X}$–valued locations Y_i's are random, and a measure with random masses at fixed locations in $\mathbb{X}$. Accordingly

$$\tilde{\mu} = \tilde{\mu}_c + \sum_{i=1}^{M} V_i \, \delta_{z_i} \tag{A.1}$$

where the fixed jump points $z_1, \ldots, z_M$, with $M \in \{1, 2, \ldots, +\infty\}$, are in $\mathbb{X}$, the (non–negative) random jumps $V_1, \ldots, V_M$ are mutually independent and they are independent from $\tilde{\mu}_c$. Finally, $\tilde{\mu}_c$ is characterized by the Laplace functional

$$\mathbb{E}\Big[\mathrm{e}^{-\int_{\mathbb{X}} f(x)\, \tilde{\mu}_c(\mathrm{d}x)}\Big] = \exp\Big\{-\int_{\mathbb{R}^+ \times \mathbb{X}} \big[1 - \mathrm{e}^{-s f(x)}\big]\, \nu(\mathrm{d}s, \mathrm{d}x)\Big\} \tag{A.2}$$

where $f : \mathbb{X} \to \mathbb{R}$ is a measurable function such that $\int |f| \, \mathrm{d}\tilde{\mu}_c < \infty$ (almost surely) and ν is a measure on $\mathbb{R}^+ \times \mathbb{X}$ such that

$$\int_B \int_{\mathbb{R}^+} \min\{s, 1\}\, \nu(\mathrm{d}s, \mathrm{d}x) < \infty \tag{A.3}$$

for any B in $\mathscr{X}$. From (A.2), which provides a Lévy-Khintchine representation of CRMs, it is apparent that they are closely connected to Poisson processes. Indeed, $\tilde{\mu}_c$ can be represented as a linear functional of a Poisson process $\tilde{M}$ on $\mathbb{R}^+ \times \mathbb{X}$ with mean measure ν. To state this precisely, $\tilde{M}$ is a random subset of $\mathbb{R}^+ \times \mathbb{X}$ and if $\tilde{N}(A) = \mathrm{card}(\tilde{M} \cap A)$ for any $A \subset \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X}$ such that $\nu(A) < \infty$, then

$$\mathbb{P}[\tilde{N}(A) = k] = \frac{(\nu(A))^k \, \mathrm{e}^{-\nu(A)}}{k!} \qquad k = 0, 1, 2, \ldots .$$

It can then be shown that

$$\tilde{\mu}_c(A) = \int_A \int_{\mathbb{R}^+} s \, \tilde{N}(\mathrm{d}s, \mathrm{d}x) \qquad \forall A \in \mathscr{X}. \tag{A.4}$$

The measure ν characterizing $\tilde{\mu}_c$ is referred to as the *Lévy or Poisson intensity* of $\tilde{\mu}_c$: it contains all the information about the distributions of the jumps and locations of $\tilde{\mu}_c$. It is often useful to separate the jump and location part of ν by writing it as

$$\nu(\mathrm{d}s, \mathrm{d}x) = \rho_x(\mathrm{d}s)\, \alpha(\mathrm{d}x) \tag{A.5}$$

where α is a measure on $(\mathbb{X}, \mathscr{X})$ and ρ a transition kernel on $\mathbb{X} \times \mathscr{B}(\mathbb{R}^+)$, *i.e.* $x \mapsto \rho_x(A)$ is $\mathscr{X}$–measurable for any A in $\mathscr{B}(\mathbb{R}^+)$ and ρ_x is a measure

on $(\mathbb{R}^+, \mathscr{B}(\mathbb{R}^+))$ for any x in $\mathbb{X}$. If $\rho_x = \rho$ for any x, then the distribution of the jumps of $\tilde{\mu}_c$ is independent of their location and both ν and $\tilde{\mu}_c$ are termed *homogeneous*. Otherwise, ν and $\tilde{\mu}_c$ are termed *non–homogeneous*.

Another important property of CRMs is their almost sure discreteness [53], which means that their realizations are discrete measures with probability 1. This fact essentially entails discreteness of random probability measures obtained as transformations of CRMs. It is worth noting that almost sure discreteness of the Dirichlet process was first shown in [3].

Finally, note that, if $\tilde{\mu}$ is defined on $\mathbb{X} = \mathbb{R}$, one can also consider the càdlàg random distribution function induced by $\tilde{\mu}$, namely $\{\tilde{\mu}((-\infty, x]) : x \in \mathbb{R}\}$. Such a random function defines an *increasing additive process*, that is a process whose increments are non–negative, independent and possibly not stationary. See [77] for an exhaustive account. We close this section by introducing three noteworthy examples of CRMs.

Example A.1. *(The gamma CRM).* A homogeneous CRM $\tilde{\gamma}$ whose Lévy intensity is given by

$$\nu(\mathrm{d}s, \mathrm{d}x) = \frac{\mathrm{e}^{-s}}{s}\,\mathrm{d}s\,\alpha(\mathrm{d}x) \tag{A.6}$$

is a *gamma process* with parameter measure α on $\mathbb{X}$. It is characterized by its Laplace functional which is given by

$$\mathbb{E}\left[\mathrm{e}^{-\int f\,\mathrm{d}\tilde{\gamma}}\right] = \mathrm{e}^{-\int \log(1+f)\,\mathrm{d}\alpha} \tag{A.7}$$

for any measurable function $f : \mathbb{X} \to \mathbb{R}$ such that $\int \log(1 + |f|)\,\mathrm{d}\alpha < \infty$. Now set $f = \lambda \mathbb{1}_B$ with $\lambda > 0$, $B \in \mathscr{X}$ such that $\alpha(B) < \infty$ and $\mathbb{1}_B$ denoting the indicator function of set B. In this case one obtains

$$\mathbb{E}\left[\mathrm{e}^{-\lambda\,\tilde{\gamma}(B)}\right] = [1 + \lambda]^{-\alpha(B)},$$

from which it is apparent that $\tilde{\gamma}(B)$ has a gamma distribution with scale and shape parameter equal to 1 and $\alpha(B)$, respectively. □

Example A.2. *(The generalized gamma CRM).* Let $\sigma \in (0, 1)$, $\tau > 0$, α be a boundedly finite measure on $\mathbb{X}$ and consider a CRM $\tilde{\mu}$ with Lévy intensity defined by

$$\nu(\mathrm{d}v, \mathrm{d}x) = \frac{\sigma}{\Gamma(1-\sigma)}\,\frac{\mathrm{e}^{-\tau v}}{v^{1+\sigma}}\,\mathrm{d}v\,\alpha(\mathrm{d}x). \tag{A.8}$$

Such CRMs [4] are known as the family of generalized gamma CRMs and can be characterized as the exponential family generated by the positive stable laws (see [4, 67]). It includes, as special cases, the σ–stable CRM for $\tau = 0$, the inverse Gaussian CRM for $\sigma = 1/2$ and the gamma CRM as $\sigma \to 0$. The generalized gamma CRM has Laplace functional of the form

$$\mathbb{E}\left[\mathrm{e}^{-\int f\,\mathrm{d}\tilde{\mu}}\right] = \mathrm{e}^{-\int (\tau+f)^{\sigma}\,\mathrm{d}\alpha - \tau^{\sigma}\alpha(\mathbb{X})} \tag{A.9}$$

for any measurable function $f : \mathbb{X} \to \mathbb{R}$ such that $\int(\tau + |f|)^{\sigma}\mathrm{d}\alpha < \infty$. For instance, if $\sigma = 1/2$ and $f = \lambda \mathbb{1}_B$ with $\lambda > 0$, $B \in \mathscr{X}$ such that $\alpha(B) < \infty$ one obtains

$$\mathbb{E}\left[\mathrm{e}^{-\lambda\,\tilde{\mu}(B)}\right] = \mathrm{e}^{-\alpha(B)\sqrt{\tau+\lambda}-\sqrt{\tau}},$$

which means that $\tilde{\mu}(B)$ has an inverse Gaussian distribution with parameters τ and $\alpha(B)$. □

Example A.3. *(The extended gamma CRM).* Another interesting CRM, introduced in [18], is the extended gamma CRM (also known as weighted gamma CRM according to the terminology of [61]). It differs from the two previous examples in that it is a non–homogeneous CRM or in other words the distribution of the jumps depends on their location. Letting β be a positive real valued function and α be a boundedly finite measure on $\mathbb{X}$, an extended gamma CRM is characterized by the Lévy intensity

$$\nu(\mathrm{d}v, \mathrm{d}x) = \frac{\mathrm{e}^{-\beta(x)v}}{v}\,\mathrm{d}v\,\alpha(\mathrm{d}x). \tag{A.10}$$

The corresponding Laplace functional is given by

$$\mathbb{E}\left[\mathrm{e}^{-\int f\,\mathrm{d}\tilde{\mu}}\right] = \mathrm{e}^{-\int \log(1+f\,\beta^{-1})\mathrm{d}\alpha} \tag{A.11}$$

for any measurable function $f : \mathbb{X} \to \mathbb{R}$ such that $\int \log(1 + |f|\,\beta^{-1})\mathrm{d}\alpha < +\infty$. Infinitesimally speaking, one has that $\tilde{\mu}(\mathrm{d}x)$ is gamma distributed with scale parameter $\beta(x)$ and shape parameter $\alpha(\mathrm{d}x)$. □

Appendix B: Transforms and inversion formulae

Most of the results presented in Sections 2.2, 3.1 and 3.2 depend on the inversion of integral transforms. Here we provide a few details on their definition and the inversion formulae one resorts to in order to obtain the results we have been describing.

B.1. Generalized Stieltjes transform

Let X be a non–negative random variable whose probability distribution is absolutely continuous with respect to the Lebesgue measure with density function g. The generalized Stieltjes transform of order $p > 0$ of X, or of the corresponding density g, is defined by

$$\mathcal{S}_p[z; X] = \int_0^{\infty} \frac{g(x)}{(x+z)^p}\,\mathrm{d}x$$

for any $z \in \mathbb{C}$ such that $|\arg(z)| < \pi$. A special case occurs when $p = 1$ so that the Stieltjes transform is, at least formally, an iterated Laplace transform

$$\mathcal{S}_1[z; X] = \int_0^{\infty} \mathrm{e}^{-uz} \int_0^{\infty} \mathrm{e}^{-ux}\, g(x)\,\mathrm{d}x\,\mathrm{d}u.$$

Moreover, the inversion formula which allows to recover g from $\mathcal{S}_1[\,\cdot\,; X]$ is simple since it yields

$$g(x) = \lim_{\epsilon \downarrow 0} \frac{1}{2\pi \mathrm{i}} \left\{ \mathcal{S}_1[-x - \mathrm{i}\epsilon; X] - \mathcal{S}_1[-x + \mathrm{i}\epsilon; X] \right\}.$$

Stieltjes transforms of order $p = 1$ are also very useful for proving infinite divisibility of probability distributions. See, e.g., [39].

The inversion formula becomes more complicated when $p \neq 1$. Different versions of it can be found in the literature. See, e.g, [82] and [79]. Here we refer to the result contained in the latter contribution. Under suitable conditions, for example $\mathcal{S}_p[\,\cdot\,; X]$ is holomorphic on $\{z \in \mathbb{C} : \, |\arg(z)| < \pi\}$ and $|z^\beta \, \mathcal{S}_p[z; X]|$ is bounded at infinity for some $\beta > 0$, one has

$$g(x) = -\frac{x^p}{2\pi \mathrm{i}} \int_\Gamma (1+w)^{p-1} \, \mathcal{S}_p'[xw; X] \, \mathrm{d}w \tag{B.1}$$

where Γ is a contour in the complex plane starting and ending at the point $w = -1$ and enclosing the origin in a counterclockwise sense. If $p < 1$, then one can integrate by parts the expression above thus obtaining

$$g(x) = \frac{p-1}{2\pi \mathrm{i}} \, x^{p-1} \int_\Gamma (1+w)^{p-2} \, \mathcal{S}_p[xw; X] \, \mathrm{d}w. \tag{B.2}$$

B.2. Characteristic functions

In order to determine prior and posterior distributions of means of NRMIs we have applied an inversion formula for characteristics functions provided in Gurland [31]. The formula is useful when one is interested in determining the probability distribution of ratios of random variables. And this is the case for means of NRMIs. If F is a distribution functions and ϕ the corresponding characteristic function, then

$$\begin{aligned} F(x) - F(x-) &= 1 - \frac{1}{\pi \mathrm{i}} \lim_{\epsilon \downarrow 0, \, T \uparrow \infty} \left(\int_{-T}^{-\epsilon} + \int_{\epsilon}^{T} \right) \mathrm{e}^{-\mathrm{i}xt} \, \frac{\phi(t)}{t} \, \mathrm{d}t \\ &= 1 - \frac{2}{\pi} \lim_{\epsilon \downarrow 0, \, T \uparrow \infty} \int_{\epsilon}^{T} \frac{1}{t} \, \mathrm{Im} \left[\mathrm{e}^{-\mathrm{i}xt} \, \phi(t) \right] \, \mathrm{d}t. \end{aligned}$$

If one sets $x = 0$, the expression in (3.14) is easily obtained.

Acknowledgements

The authors are grateful to the Associate Editor and an anonymous referee for their valuable comments and suggestions that greatly improved the presentation of the paper. A. Lijoi and I Prünster are partially supported by MIUR–Italy, project "Bayesian methods: theoretical developments and novel applications".

References

[1] Barlow, M., Pitman, J. and Yor, M. (1989). Une extension multidimensionnelle de la loi de l'arc sinus. In *Séminaire de Probabilités XXIII* (Azema, J., Meyer, P.-A. and Yor, M., Eds.), 294–314, Lecture Notes in Mathematics **1372**. Springer, Berlin. MR1022893

[2] Bertoin, J. and Yor, M. (2005). Exponential functionals of Lévy processes. *Prob. Surveys* **2**, 191–212. MR2178044

[3] Blackwell, D. (1973). Discreteness of Ferguson selections. *Ann. Statist.* **1**, 356-358. MR0348905

[4] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. Appl. Probab.* **31**, 929–953. MR1747450

[5] Carlton, M.A. (2002). A family of densities derived from the three-parameter Dirichlet process. *J. Appl. Probab.* **39**, 764–774. MR1938169

[6] Carmona, P., Petit, F. and Yor, M. (1997). On the distribution and asymptotic results for exponential functionals of Lévy processes. In *Exponential functionals and principal values related to Brownian motion* (M. Yor, ed.), 73–126. Biblioteca de la Revista Matematica Ibero–Americana, Madrid. MR1648657

[7] Cifarelli, D.M. and Melilli, E. (2000). Some new results for Dirichlet priors. *Ann. Statist.* **28**, 1390–1413. MR1805789

[8] Cifarelli, D.M. and Regazzini, E. (1979a). A general approach to Bayesian analysis of nonparametric problems. The associative mean values within the framework of the Dirichlet process. I. (Italian) *Riv. Mat. Sci. Econom. Social.* **2**, 39–52. MR0573686

[9] Cifarelli, D.M. and Regazzini, E. (1979b). A general approach to Bayesian analysis of nonparametric problems. The associative mean values within the framework of the Dirichlet process. II. (Italian) *Riv. Mat. Sci. Econom. Social.* **2**, 95–111. MR0574040

[10] Cifarelli, D.M. and Regazzini, E. (1990). Distribution functions of means of a Dirichlet process. *Ann. Statist.*, **18**, 429–442 (Correction in *Ann. Statist.* (1994) **22**, 1633–1634). MR1041402

[11] Cifarelli, D.M. and Regazzini, E. (1993). Some remarks on the distribution functions of means of a Dirichlet process. *Technical Report* **93.4**, IMATI–CNR, Milano.

[12] Cifarelli, D.M. and Regazzini, E. (1996). Tail-behaviour and finiteness of means of distributions chosen from a Dirichlet process. *Technical Report* **96.19**, IMATI–CNR, Milano. Available at: http://www.mi.imati.cnr.it/iami/papers/96-19.ps

[13] Constantine, G.M. and Savits, T.H. (1996). A multivariate Faà di Bruno formula with applications. *Trans. Amer. Math. Soc.* **348**, 503–520. MR1325915

[14] Daley, D.J. and Vere–Jones, D. (1988). *An introduction to the theory of point processes.* Springer, New York. MR0950166

[15] Diaconis, P. and Kemperman, J. (1996). Some new tools on Dirichlet priors. In *Bayesian statistics 5* (J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, eds.) 97–106. Oxford Univ. Press. MR1425401
[16] Dickey, J.M. and Jiang, T.J. (1998). Filtered-variate prior distributions for histogram smoothing. *J. Amer. Statist. Assoc.* **93**, 651–662. MR1631349
[17] Doksum, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.* **2**, 183–201. MR0373081
[18] Dykstra, R.L. and Laud, P. (1981). A Bayesian nonparametric approach to reliability. *Ann. Statist.* **9**, 356–367. MR0606619
[19] Epifani, I., Guglielmi, A. and Melilli, E. (2006). A stochastic equation for the law of the random Dirichlet variance. *Statist. Probab. Lett.* **76**, 495–502. MR2266603
[20] Epifani, I., Lijoi, A. and Prünster, I. (2003). Exponential functionals and means of neutral–to–the–right priors. *Biometrika* **90**, 791–808. MR2024758
[21] Erhardsson, T. (2008). Nonparametric Bayesian inference for integrals with respect to an unknown finite Mmeasure. *Scand. J. Statist.* **35**, 354–368. MR2418747
[22] Evans, M.R. and Hanney, T. (2005). Nonequilibrium statistical mechanics of the zero-range process and related models. *J. Phys. A* **38**, R195-R240. MR2145800
[23] Exton, H. (1976). *Multiple hypergeometric functions and applications.* Ellis Horwood, Chichester. MR0422713
[24] Favaro, S., Guglielmi, A. and Walker, S.G. (2009). A class of measure-valued Markov Chains and Bayesian Nonparametrics. *Tech. Report.*
[25] Feigin, P.D. and Tweedie, R.L. (1989). Linear functionals and Markov chains associated with Dirichlet processes. *Math. Proc. Camb. Phil. Soc.* **105**, 579–585. MR0985694
[26] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209–230. MR0350949
[27] Ferguson, T.S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2**, 615–629. MR0438568
[28] Guglielmi, A. (1998). A simple procedure calculating the generalized Stieltjes transform of the mean of a Dirichlet process. *Statist. Probab. Lett.* **38**, 299–303. MR1631205
[29] Guglielmi, A. and Tweedie, R.L. (2001). Markov chain Monte Carlo estimation of the law of the mean of a Dirichlet process. *Bernoulli* **7**, 573–592. MR1849368
[30] Guglielmi, A., Holmes, C.C. and Walker, S.G. (2002). Perfect simulation involving functionals of a Dirichlet process. *J. Comput. Graph. Statist.* **11**, 306–310. MR1938137
[31] Gurland, J. (1948). Inversion formulae for the distributions of ratios. *Ann. Math. Statist.*, **19**, 228–237. MR0025005

[32] Haas, B., Miermont, G., Pitman, J. and Winkel, M. (2008). Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *Ann. Probab.* **36**, 1790–1837. MR2440924

[33] Hannum, R.C., Hollander, M. and Langberg, N.A. (1981). Distributional results for random functionals of a Dirichlet process. *Ann. Probab.* **9**, 665–670. MR0630318

[34] Hill, T. and Monticino, M. (1998). Constructions of random distributions via sequential barycenters. *Ann. Statist.* **26**, 1242–1253. MR1647649

[35] Hjort, N.L. (2003). Topics in non-parametric Bayesian statistics. In: *Highly structured stochastic systems* (P. Green, N.L. Hjort and S. Richardson, Eds.), 455–487. Oxford Univ. Press, Oxford. MR2082419

[36] Hjort, N.L. and Ongaro, A. (2005). Exact inference for random Dirichlet means. *Stat. Inference Stoch. Process.* **8**, 227–254. MR2177313

[37] Hjort, N.L. and Petrone, S. (2007). Nonparametric quantile inference using Dirichlet processes. In *Advances in statistical modeling and inference*, 463–492, World Sci. Publ., Hackensack, NJ. MR2416129

[38] Ishwaran, H. and James, L.F. (2003). Generalized weighted Chinese restaurant processes for species sampling mixture models. *Statist. Sinica* **13**, 1211–1235. MR2026070

[39] Ismail, M.E.H. and Kelker, D.H. (1979). Special functions, Stieltjes transforms and infinite divisibility. *SIAM J. Math. Anal.* **10**, 884–901. MR0541088

[40] James, L.F. (2002). Poisson process partition calculus with applications to exchangeable models and Bayesian nonparametrics. *Mathematics ArXiv*, math.PR/0205093.

[41] James, L.F. (2006). Functionals of Dirichlet processes, the Cifarelli-Regazzini identity and beta-gamma processes. *Ann. Statist.* **33**, 647–660. MR2163155

[42] James, L.F. (2007). New Dirichlet mean identities. *Mathematics Arxiv*, math.PR/0708.0614.

[43] James, L.F., Lijoi, A. and Prünster, I. (2006). Conjugacy as a distinctive feature of the Dirichlet process. *Scand. J. Statist.* **33**, 105–120. MR2255112

[44] James, L.F., Lijoi, A. and Prünster, I. (2008). Distributions of linear functionals of two parameter Poisson–Dirichlet random measures. *Ann. Appl. Probab.* **18**, 521–551. MR2398765

[45] James, L.F., Lijoi, A. and Prünster, I. (2009). Posterior analysis for normalized random measures with independent increments. *Scand. J. Statist.* **36**, 76–97. MR2508332

[46] James, L.F., Lijoi, A. and Prünster, I. (2009). On the posterior distribution of classes of random means. *Bernoulli*, DOI: 10.3150/09-BEJ200.

[47] James, L.F., Roynette, B. and Yor, M. (2008). Generalized gamma convolutions, Dirichlet means, Thorin measures, with explicit examples. *Probab. Surv.* **5**, 346–415. MR2476736

[48] Jarner, S.F., Tweedie, R.L. (2002). Convergence rates and moments of Markov chains associated with the mean of Dirichlet processes.*Stochastic Process. Appl.* **101**, 257–271. MR1931269
[49] Kerov, S.V. (1998). Interlacing measures. In *Kirillov's seminar on representation theory.* Amer. Math. Soc. Transl. Ser. 2, Vol. **181**. Amer. Math. Soc., Providence, 35–83. MR1618739
[50] Kerov, S.V. and Tsilevich, N.V. (2001). The Markov-Krein correspondence in several dimensions. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **283**, 98–122. (English translation in *J. Math. Sci.* **121**, 2345–2359). MR1879065
[51] Kingman, J.F.C. (1967). Completely random measures. *Pacific J. Math.* **21**, 59–78. MR0210185
[52] Kingman, J.F.C. (1975). Random discrete distributions (with discussion). *J. Roy. Statist. Soc. Ser. B* **37**, 1–22. MR0368264
[53] Kingman, J.F.C. (1993). *Poisson processes.* Oxford University Press, Oxford. MR1207584
[54] Lamperti, J. (1958). An occupation time theorem for a class of stochastic processes. *Trans. Amer. Math. Soc.* **88**, 380–387. MR0094863
[55] Lévy, P. (1939). Sur certains processus stochastiques homogènes. *Compositio Math.* **7**, 283–339. MR0000919
[56] Lijoi, A., Mena, R.H. and Prünster, I. (2005). Hierarchical mixture modelling with normalized inverse Gaussian priors. *J. Amer. Statist. Assoc.* **100**, 1278–1291. MR2236441
[57] Lijoi, A., Mena, R.H. and Prünster, I. (2005). Bayesian nonparametric analysis for a generalized Dirichlet process prior. *Statist. Inference Stoch. Process.*, **8**, 283–309. MR2177315
[58] Lijoi, A., Mena, R.H. and Prünster, I. (2007). Controlling the reinforcement in Bayesian nonparametric mixture models. *J. Roy. Statist. Soc. Ser. B* **69**, 715–740. MR2370077
[59] Lijoi, A. and Regazzini, E. (2004). Means of a Dirichlet process and multiple hypergeometric functions. *Ann. Probab.* **32**, 1469–1495. MR2060305
[60] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12**, 351–357. MR0733519
[61] Lo, A.Y. and Weng, C.-S. (1989). On a class of Bayesian nonparametric estimates. II. Hazard rate estimates. *Ann. Inst. Statist. Math.* **41**, 227–245. MR1006487
[62] Muliere, P. and Tardella, L. (1998). Approximating distributions of random functionals of Ferguson-Dirichlet priors. *Canad. J. Statist.* **26** 283–297. MR1648431
[63] Nieto-Barajas, L.E., Prünster, I. and Walker, S.G. (2004). Normalized random measures driven by increasing additive processes. *Ann. Statist.* **32**, 2343–2360. MR2153987
[64] Peccati, G. (2008). Multiple integral representation for functionals of Dirichlet processes. *Bernoulli* **14**, 91–124. MR2401655
[65] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields.* **102**, 145–158. MR1337249

[66] Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory* (T.S. Ferguson, L.S. Shapley and J.B. MacQueen, Eds.). IMS Lecture Notes Monogr. Ser., Vol. **30**. Inst. Math. Statist., Hayward, 245–267. MR1481784

[67] Pitman, J. (2003). Poisson-Kingman partitions. In *Statistics and science: a Festschrift for Terry Speed* (D.R Goldstein, Ed.). IMS Lecture Notes Monogr. Ser., Vol. **40**. Inst. Math. Statist., Beachwood, 1–34. MR2004330

[68] Pitman, J. and Yor, M. (1992). Arcsine laws and interval partitions derived from a stable subordinator. *Proc. London Math. Soc.* **65** 326–356. MR1168191

[69] Pitman, J., and Yor, M. (1997). On the relative lengths of excursions derived from a stable subordinator. In: *Séminaire de Probabilités XXXI* (Azema, J., Emery, M. and Yor, M., Eds.), 287–305, Lecture Notes in Mathematics **1655**. Springer, Berlin. MR1478738

[70] Pulkkinen, O. (2007). Boundary driven zero-range processes in random media. *J. Stat. Phys.* **128**, 1289–1305. MR2352010

[71] Regazzini, E. (1998). An example of the interplay between statistics and special functions. In *Atti dei convegni Lincei. Tricomi's ideas and contemporary applied mathematics* **147**, 303–320. MR1737502

[72] Regazzini, E., Guglielmi, A. and Di Nunno, G. (2002). Theory and numerical analysis for exact distribution of functionals of a Dirichlet process. *Ann. Statist.* **30**, 1376–1411. MR1936323

[73] Regazzini, E., Lijoi, A. and Prünster, I. (2003). Distributional results for means of random measures with independent increments. *Ann. Statist.* **31**, 560–585. MR1983542

[74] Regazzini, E. and Sazonov, V.V. (2001). Approximation of laws of random probabilities by mixtures of Dirichlet distributions with applications to nonparametric Bayesian inference. *Th. Prob. Appl.*, **45**, 93–110. MR1810976

[75] Romik, D. (2004). Explicit formulas for hook walks on continual Young diagrams. *Adv. in Appl. Math.* **32**, 625–654. MR2053837

[76] Romik, D. (2005). Roots of the derivative of a polynomial. *Amer. Math. Monthly* **112**, 66–69. MR2110114

[77] Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press, Cambridge. MR1739520

[78] Sethuraman, J. (1994). A constructive definition of the Dirichlet process prior. *Statist. Sinica* **2**, 639–650. MR1309433

[79] Schwarz, J.H. (2005) The generalized Stieltjes transform and its inverse. *J. Math. Phys.* **46** 013501-8. MR2113770

[80] Spitzer, F. (1970). Interaction of Markov processes. *Advances in Math.* **5**, 246-290. MR0268959

[81] Stanley, R.P. (1999). *Enumerative combinatorics. Volume 2.* Cambridge University Press. MR1676282

[82] Sumner, D.B. (1949). An inversion formula for the generalized Stieltjes transform. *Bull. Amer. Math. Soc.* **55**, 174–183. MR0027888

[83] Tamura, H. (1988). Estimation of rare errors using expert judgement. *Biometrika* **75**, 1–9. MR0932811

[84] Tsilevich, N.V. (1997). Distribution of mean values for some random measures. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **240**, 268–279. (English translation in *J. Math. Sci.* **96**, 3616–3623). MR1691650

[85] Vershik, A.M., Yor, M. and Tsilevich, N.V. (2001). The Markov-Krein identity and the quasi-invariance of the gamma process. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **283**, 21–36. (English translation in *J. Math. Sci.* **121**, 2303–2310). MR1879060

[86] Walker, S.G. and Muliere, P. (1997). Beta-Stacy processes and a generalization of the Pólya-urn scheme. *Ann. Statist.* **25**, 1762–1780. MR1463574

[87] Yamato, H. (1980). On behaviors of means of distributions with Dirichlet processes. *Rep. Fac. Sci. Kagoshima Univ.* **13**, 41–45. MR0618622

[88] Yamato, H. (1984a). Expectations of functions of samples from distributions chosen from Dirichlet processes. *Rep. Fac. Sci. Kagoshima Univ. Math. Phys. Chem.* **17**, 1–8. MR0798590

[89] Yamato, H. (1984b). Characteristic functions of means of distributions chosen from a Dirichlet process. *Ann. Probab.* **12**, 262–267. MR0723745

[90] Yano, K. and Yano, Y. (2008). Remarks on the density of the law of the occupation time for Bessel bridges and stable excursions. *Statist. Prob. Letters* **78**, 2175-2180. MR2458027