Electronic Journal of Statistics
Vol. 7 (2013) 62–90
ISSN: 1935-7524
DOI: 10.1214/12-EJS764

A vector of Dirichlet processes

Fabrizio Leisen

Universidad Carlos III de Madrid
Departamento de Estadistica
Calle Madrid 126, 28903 Getafe (Madrid), Spain
e-mail: fabrizio.leisen@gmail.com

Antonio Lijoi*

University of Pavia
Department of Economics and Management
Via San Felice 5, 27100 Pavia, Italy
e-mail: lijoi@unipv.it

and

Dario Spanó

University of Warwick
Department of Statistics
Coventry CV4 7AL, United Kingdom
e-mail: d.spano@warwick.ac.uk

Abstract: Random probability vectors are of great interest especially in view of their application to statistical inference. Indeed, they can be used for identifying the de Finetti mixing measure in the representation of the law of a partially exchangeable array of random elements taking values in a separable and complete metric space. In this paper we describe the construction of a vector of Dirichlet processes based on the normalization of an exchangeable vector of completely random measures that are jointly infinitely divisible. After deducing the form of the multivariate Laplace exponent associated to the vector of the gamma completely random measures, we analyze some of their distributional properties. Our attention particularly focuses on the dependence structure and the specific partition probability function induced by the proposed vector.

Keywords and phrases: Bayesian nonparametrics, completely random measure, Dirichlet process, Lévy copula, multivariate Lévy measure, partial exchangeability, partition probability function.

Received August 2012.

1. Introduction

Random probability measures represent the key ingredient for the actual implementation of Bayesian nonparametric procedures and the Dirichlet process is the most celebrated example. Indeed, their probability distribution plays the

*Also affiliated to Collegio Carlo Alberto, via Real Collegio 30, 10024 Moncalieri (TO), Italy.

role of a prior: this is combined with the data to yield a posterior, or a predictive, distribution that is used to determine either exact or approximate Bayesian inferences on quantities of interest. The most notable advances in the field have been achieved under the assumption that the observations $X_1, \dots, X_N$ are the initial segment of an (ideally) infinite *exchangeable* sequence of random elements $(X_i)_{i \geq 1}$ taking values in some complete and separable metric space $\mathbb{X}$ endowed with the Borel σ–algebra $\mathscr{X}$. If $\boldsymbol{P}_{\mathbb{X}}$ stands for the set of all probability measures on $(\mathbb{X}, \mathscr{X})$ and $\mathscr{P}_{\mathbb{X}}$ is the σ–algebra induced by the topology of weak convergence on $\boldsymbol{P}_{\mathbb{X}}$, the well–known de Finetti representation theorem leads to a mixture representation of the finite–dimensional distributions of $(X_i)_{i \geq 1}$ of the form

$$\mathbb{P}[(X_1, \dots, X_N) \in A] = \int_{\boldsymbol{P}_{\mathbb{X}}} p^N(A)\, Q(\mathrm{d}p) \qquad \forall A \in \mathscr{X}^N, \quad \forall N \geq 1 \qquad (1)$$

where $p^N = \prod_1^N p$ and Q is a probability measure on $(\boldsymbol{P}_{\mathbb{X}}, \mathscr{P}_{\mathbb{X}})$ that takes on the name of *de Finetti* measure of the sequence $(X_i)_{i \geq 1}$. This is equivalent to stating that the X_i's are conditionally independent, namely $\mathbb{P}[X_1 \in A_1, \dots, X_N \in A_N \,|\, \tilde{p}] = \prod_{i=1}^N \tilde{p}(A_i)$ with $\tilde{p} \sim Q$.

One of the most popular applications of nonparametric priors to statistical inference concerns density estimation and is based on hierarchical mixture models where the X_i's are latent variables and Q is the law of a Dirichlet process. See [21]. The dramatic advances in the implementation of Markov Chain Monte Carlo simulation algorithms in the last two decades have, then, made Bayesian nonparametric methods directly applicable to a wide range of real world problems. Moreover, a considerable body of work has been devoted to the proposal of alternatives to the Dirichlet process, i.e. different Q's in (1), for modelling exchangeable random elements. See [20] for a recent review.

1.1. Partial exchangeability

The notion of exchangeability characterized through (1) corresponds to a more intuitive idea of homogeneity among the observations. While this seems to be a natural assumption if one aims at prediction, it does not typically correspond to many situations of practical interest where data are originated under different experimental conditions or are, more generally, indexed by a collection of covariates that are relevant for statistical inference. This is the case of data related to different studies for which it is reasonable to assume exchangeability within and not between studies. For example, when comparing the survival time of two groups of patients displaying the same pathology and treated with two different therapies, the data can be thought of as labeled by a binary covariate identifying the specific treatment a patient is subject to. In this case, it is not realistic to assume that any two observations labeled by distinct covariate values, i.e. originating from different groups, are "homogeneous". At the same time, one might be interested in clustering patients subject to a therapy and borrow information from the outcomes on patients in the other group. Similar issues arise in more general regression problems where data are labeled by a collection of covariates taking values in a general space, and exchangeability can be assumed only for observations associated to the same covariate realization.

There clearly is a wide range of applied problems which motivate the great interest that has recently emerged for the definition of nonparametric priors accommodating for forms of dependence more general than exchangeability and able to capture the "heterogeneity" featured by the data. Here we shall focus on a proposal that is suited for handling inferential problems involving *partially exchangeable* data. In this setting, if z denotes a set of covariates taking values in $\mathscr{Z}$, the observations are elements from a collection of sequences $\{(X_i(z))_{i\geq 1} : z \in \mathscr{Z}\}$ such that for any $n \geq 1$, positive integers $q_1, \ldots, q_n$

$$\begin{aligned}\mathbb{P}[(\boldsymbol{X}^{q_1}(z_1), \ldots, \boldsymbol{X}^{q_n}(z_n)) \in A] \\ = \int_{\boldsymbol{P}_{\mathbb{X}}^{|\boldsymbol{q}|}} (p_{z_1}^{q_1} \times \cdots \times p_{z_n}^{q_n})(A)\, Q_{\mathscr{Z}}(\mathrm{d}p_{z_1}, \ldots, \mathrm{d}p_{z_n})\end{aligned} \tag{2}$$

for any $A \in \mathscr{X}^{|\boldsymbol{q}|}$, where $\boldsymbol{X}^{q_j}(z_j) = (X_1(z_j), \ldots, X_{q_j}(z_j))$, $\boldsymbol{q} = (q_1, \ldots, q_n)$ and $|\boldsymbol{q}| = q_1 + \cdots + q_n$. Here $Q_{\mathscr{Z}}$ is the probability distribution of a collection of possibly dependent random probability measures $\tilde{\boldsymbol{p}} := \{\tilde{p}_z : z \in \mathscr{Z}\}$ and plays the role of prior distribution in a similar fashion as Q does in the exchangeable case (1). Indeed, note that the mixture representation in (2) is due to de Finetti in [5] and it amounts to assuming exchangeability for the sequence of observations $(X_i(z))_{i\geq 1}$ corresponding to the same covariate value z in $\mathscr{Z}$ and conditional independence among observations corresponding to different covariate values. Hence, data corresponding to different covariate realizations are not conditionally identically distributed. The model can be equivalently summarized in a hierarchical form as follows

$$\begin{aligned} X_i(z_j) \,|\, \tilde{\boldsymbol{p}} &\overset{\text{iid}}{\sim} \tilde{p}_{z_j} && i = 1, \ldots, q_{n_j} \\ (X_i(z_j), X_\ell(z_\kappa)) \,|\, \tilde{\boldsymbol{p}} &\overset{\text{iid}}{\sim} \tilde{p}_{z_j} \times \tilde{p}_{z_\kappa} && i = 1, \ldots, q_{n_j};\ \ \ell = 1, \ldots, q_{n_\kappa} \\ \tilde{\boldsymbol{p}} &\sim Q_{\mathscr{Z}} \end{aligned}$$

When $Q_{\mathscr{Z}}$ is degenerate on $\boldsymbol{P}_{\mathbb{X}}$ so that $\tilde{p}_z = \tilde{p}_{z'}$ (almost surely) for any $z \neq z'$, there is no heterogeneity among the data and exchangeability is recovered.

To our knowledge, the first proposal of a nonparametric prior in this framework has been provided in [2], where $\mathscr{Z}$ is a finite set and $Q_{\mathscr{Z}}$ is the distribution of a mixture of products of Dirichlet processes. However, the most recent literature has been spurred by S.N. MacEachern's seminal papers [22, 23] where a dependent Dirichlet process is defined by means of a series representation. The main idea is to rely on the stick–breaking construction of the Dirichlet process and define, for each z in $\mathscr{Z}$,

$$\tilde{p}_z = \sum_{i\geq 1} \omega_{i,z}\, \delta_{Y_{i,z}} \tag{3}$$

where $\omega_{1,z} = V_{1,z}$, $\omega_{i,z} = V_{i,z} \prod_{j=1}^{i-1}(1 - V_{j,z})$, for $i \geq 2$, and $V_{i,z} \stackrel{\text{iid}}{\sim} \text{Beta}(1, \theta_z)$ for $i \geq 1$. Moreover, $(Y_{i,z})_{i\geq1}$ are independent and identically distributed (i.i.d.) from some probability distribution $P_{0,z}$. For any z, $\tilde{p}_z$ is a Dirichlet process with base measure $\alpha_z = \theta_z P_{0,z}$ and dependence between any two random probabilities $\tilde{p}_z$ and $p_{z'}$, for $z \neq z'$, is induced by possible dependence between $V_{i,z}$ and $V_{i,z'}$ and/or between $Y_{i,z}$ and $Y_{i,z'}$. Later developments that build on MacEachern's idea can be found, e.g., in [6, 12, 28, 7, 8]. A more complete and stimulating picture of the recent state of the art in the field, including a variety of interesting applications, is provided in [13]. The practical use of these models has been eased by the developments of suitable MCMC sampling algorithms, that make use of the stick–breaking representation of $\tilde{p}_z$ in (3) and lead to an approximate evaluation of Bayesian inferences. As a matter of fact, the availability of an R package, named `DPpackage`, that allows for an automatic implementation of MCMC samplers with dependent processes, has made these models accessible to interested practitioners as well. See [15] for details.

The present paper relies on a different approach and defines each dependent random probability measure $\tilde{p}_z$ by normalizing a completely random measure (CRM) defined by

$$\tilde{\mu}_z = \sum_{i\geq1} J_{i,z}\, \delta_{Y_{i,z}} \tag{4}$$

where $(J_{i,z})_{i\geq1}$ is a sequence of non–negative independent random variables such that $\sum_{i\geq1} J_{i,z} < \infty$, almost surely, and $(Y_{i,z})_{i\geq1}$ is a sequence of i.i.d random variables taking values in $\mathbb{X}$. Hence, in (3) one defines $\omega_{i,z} = J_{i,z}/\sum_{i\geq1} J_{i,z}$. In view of this definition, dependence between $\tilde{p}_z$ and $\tilde{p}_{z'}$ will be induced by the dependence between $\tilde{\mu}_z$ and $\tilde{\mu}_{z'}$ and this is summarized through a multivariate Lévy intensity. In our proposal $(Y_{i,z})_{i\geq1} = (Y_i)_{i\geq1}$ and more importantly $\tilde{\mu}_z$ is, for any $z \in \mathcal{Z}$, a gamma CRM which yields, through normalization, a Dirichlet process. Hence, the model discussed in this paper can be seen as a dependent Dirichlet process as in (3), with $\mathcal{Z}$ being a finite set. However, we are not able to establish which form of dependence among the stick–breaking weights $V_{i,z}$ leads to the dependent Dirichlet process prior that will be discussed in the next sections.

The use of dependent CRMs in the construction of dependent random probability measures has been considered in other work, as well. For example, in [9] and [17] dependence is induced through Lévy copulas that allow to define multivariate Lévy intensities with fixed marginals, thus operating in a similar fashion as traditional copulas do for probability distributions. More specifically, in [9] a class of dependent neutral to the right processes for the analysis of partially exchangeable survival data is constructed, whereas in [17] a bivariate vector of two–parameter Poisson–Dirichlet processes is introduced. Note that the latter is closer in spirit to the present work, the main distinctive feature being that it relies on the normalization of a random measure which is not completely random. As to our knowledge, before these two contributions the use of Lévy copulas has been mostly confined to applications in Finance as well detailed, e.g., in [3]. Furthermore, a construction that does not rely on Lévy copulas can be found in [19]: here the authors define dependent CRMs by means of the superposition of independent completely random measures some of which are shared. See also [26] for a similar approach aimed at applications to spatial statistics. An allied contribution that, similarly to ours, applies to the case where $\mathcal{Z}$ is a finite set is in [24] where the authors model each $\tilde{p}_z$ as a convex linear combination of Dirichlet processes one of which is shared through different z's thus inducing dependence.

1.2. Main results and outline of the paper

As already anticipated, the present paper introduces a new vector of gamma CRMs $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}_1, \ldots, \tilde{\mu}_n)$, which is characterized in terms of a Lévy intensity ν on $(\mathbb{R}^+)^n \times \mathbb{X}$. Each $\tilde{\mu}_i$ has the same marginal distribution and is such that $\tilde{\mu}_i(A)$ is gamma distributed with

$$\mathbb{E}\left[\mathrm{e}^{-\lambda\tilde{\mu}_i(A)}\right] = (1+\lambda)^{-cP_0(A)} \qquad \forall\lambda > 0, \tag{5}$$

for probability measure P_0 on $\mathbb{X}$, for any set A in $\mathscr{X}$ such that $P_0(A) > 0$ and for some constant $c > 0$. This construction has two merits. On one side it ensures mutual independence of vectors $(\tilde{\mu}_1(A_i), \ldots, \tilde{\mu}_n(A_i))$, as $i = 1, \ldots, k$, for any collection of pairwise disjoint sets $A_1, \ldots, A_k$ in $\mathscr{X}$ and for any $k \geq 1$. On the other, it allows one to evaluate exactly the Laplace transform of the vector $(\tilde{\mu}_1, \ldots, \tilde{\mu}_n)$ through

$$\mathbb{E}\left[\mathrm{e}^{-\langle\boldsymbol{\lambda},\tilde{\boldsymbol{\mu}}(A)\rangle}\right] = \exp\left\{-\int_{(\mathbb{R}^+)^n\times A}\left[1-\mathrm{e}^{-\langle\boldsymbol{\lambda},\boldsymbol{y}\rangle}\right]\,\nu(\mathrm{d}\boldsymbol{y},\mathrm{d}x)\right\} \tag{6}$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in (0,\infty)^n$, A is any set in $\mathscr{X}$, $\tilde{\boldsymbol{\mu}}(A) = (\tilde{\mu}_1(A), \ldots, \tilde{\mu}_n(A))$ and $\langle \boldsymbol{s}, \boldsymbol{t}\rangle = \sum_{i=1}^n s_i t_i$ for any $\boldsymbol{s}$ and $\boldsymbol{t}$ in $\mathbb{R}^n$. Moreover, ν is such that

$$\int_{(\mathbb{R}^+)^{n-1}} \nu(\mathrm{d}\boldsymbol{y}_{-i}, \mathrm{d}x) = \frac{\mathrm{e}^{-y_i}}{y_i}\, cP_0(\mathrm{d}x) \tag{7}$$

with $\mathrm{d}\boldsymbol{y}_{-i}$ denoting integration with respect to all $\boldsymbol{y}$ coordinates but the i–th. In the special case where $\boldsymbol{\lambda}$ is such that $\lambda_1 \neq \cdots \neq \lambda_n$ we show that

$$\mathbb{E}\left[\mathrm{e}^{-\langle\boldsymbol{\lambda},\tilde{\boldsymbol{\mu}}(A)\rangle}\right] = \exp\left\{-cP_0(A)\sum_{i=1}^n \frac{\lambda_i^{n-1}\,\log(1+\lambda_i)}{\prod_{\substack{j=1\\ j\neq i}}^n(\lambda_i-\lambda_j)}\right\}. \tag{8}$$

A more elaborated expression, still available in closed form, is recovered if any two of the λ_i's coincide. This will be detailed in one of the next sections. In particular, if $\lambda_j = 0$ for any $j \neq i$ and $\lambda_i > 0$, one obtains (5) so that $\tilde{\mu}_i$ is, indeed, a gamma completely random measure whose intensity is displayed on the right hand side of (7). Notice that the expression (8) is invariant with respect to permutations of the λ_i's, a property which remains valid even for the general case allowing for ties in $(\lambda_1, \ldots, \lambda_n)$ as it can be noted from (25). Therefore the vector $\tilde{\boldsymbol{\mu}}$ is exchangeable. This is *per se* an interesting distinctive feature since the model still preserves analytical tractability despite a non-Markovian form of dependence.

Both the availability of the Laplace transform and the independence of the increments of $\tilde{\boldsymbol{\mu}}$ are key tools for determining, e.g., the mixed moments of $\tilde{\boldsymbol{\mu}}$. If we confine here to the simple case where $n = 2$, we show that for any set A in $\mathscr{X}$

$$\mathbb{E}\,[\tilde{\mu}_1^{q_1}(A)\,\tilde{\mu}_2^{q_2}(A)] = q_1!q_2! \sum_{k=1}^{q_1+q_2} c^k P_0^k(A) \sum_{j=1}^{q_1+q_2} \sum_{(*)} \prod_{i=1}^{j} \frac{1}{\lambda_i!(s_{1,i}+s_{2,i})!} \tag{9}$$

where $(*)$ denotes the set where the vectors $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_j)$ and $\boldsymbol{s}_i = (s_{1,i}, s_{2,i})$, for $i = 1, \ldots, j$, vary in the sum. This result, that can be extended to the case of gamma CRM vectors of dimension $n > 2$, is relevant for analyzing some distributional properties useful for Bayesian nonparametric inference. Indeed, if one is willing to analyze data for which partial exchangeability is a suitable form of dependence, one can set $Q_{\mathcal{Z}}$ as the probability distribution of $(\tilde{p}_1, \ldots, \tilde{p}_n)$ where $\tilde{p}_i = \tilde{\mu}_i/\tilde{\mu}_i(\mathbb{X})$, for $i = 1, \ldots, n$. This induces a vector of dependent Dirichlet processes, i.e. a vector such that each $\tilde{p}_i$ is a Dirichlet process with base measure cP_0. Moreover, $\mathcal{Z} = \{1, \ldots, n\}$. Making use of the same technique that leads to proving (9) one can evaluate the correlation coefficient between $\tilde{p}_i(A)$ and $\tilde{p}_j(B)$, for any $i \neq j$ and $A, B \in \mathscr{X}$. In particular, it is shown that for any A in $\mathscr{X}$

$$\mathrm{corr}(\tilde{p}_1(A), \tilde{p}_2(A)) = K_c \int_0^1 \xi_c(z)\,\mathrm{d}z \tag{10}$$

for some constant K_c that depends on the total mass $c > 0$, for some function ξ_c that can be represented in terms of hypergeometric functions ${}_1F_1$ and ${}_2F_1$. The integral in (10) can be evaluated numerically and, more importantly, the right-hand side is independent from set A: this fact typically motivates the use of (10) as a measure of overall dependence between $\tilde{p}_1$ and $\tilde{p}_2$.

The illustration of the properties of $(\tilde{p}_1, \ldots, \tilde{p}_n)$ is completed by the determination of the associated partition probability function. Indeed, notice that each $\tilde{p}_i$ is discrete with probability 1 so that ties may occur among the data: this induces a partition with any two observations belonging to the same cluster if and only if they coincide. The analysis of the random partition associated to a random discrete distribution is important for computational purposes. For example, when the observations are exchangeable as in (1), the so–called exchangeable partition probability function (EPPF) is a key tool for devising a MCMC sampler for density estimation in mixture models. See, e.g., [20]. This also motivates a similar investigation in the partially exchangeable case as emphasized in [19]. We shall be able to determine a closed form expression of the partially exchangeable partition probability function (pEPPF) associated to $(\tilde{p}_1, \tilde{p}_2)$. The extension to the case $n > 2$, though feasible analytically, leads to complicated expressions that are not displayed here.

The outline of the paper is as follows. In Section 2 we shall introduce some basic facts on CRMs that are necessary for an understanding of the main results we have obtained. In Section 3 some known examples of multivariate vectors of CRMs are provided and a new vector of dependent gamma CRMs is introduced. Section 4 is devoted to a detailed description of Lévy copulas and of their connection with our vector of dependent gamma CRMs. As already pointed out, attention will be focused in Section 5 on the determination of the Laplace transform of the vector since it is a key quantity for the evaluation of posterior inferences in a Bayesian nonparametric framework. Finally, a normalization procedure is adopted to yield a collection of dependent Dirichlet processes which correspond to a covariate space $\mathcal{Z} = \{1, \ldots, n\}$, for some positive integer n: we shall examine their dependence structure and state a result concerning the partially exchangeable partition probability function in Section 6.

2. Normalized completely random measures

Among the different possibilities emerged in the literature for defining a probability distribution Q on $(\boldsymbol{P}_{\mathbb{X}}, \mathscr{P}_{\mathbb{X}})$ in (1), we shall focus on a strategy that makes use of completely random measures. For this reason we devote the present section to the introduction of some preliminary material, which also serves for clarifying the main notation that will be used henceforth.

Denote by $\boldsymbol{M}_{\mathbb{X}}$ that space of boundedly finite measures on $(\mathbb{X}, \mathscr{X})$, namely each element μ in $\boldsymbol{M}_{\mathbb{X}}$ is a measure on $(\mathbb{X}, \mathscr{X})$ such that $\mu(B) < \infty$ for any bounded set $B \in \mathscr{X}$. The space $\boldsymbol{M}_{\mathbb{X}}$ is endowed with the so–called weak$^\sharp$ topology (see [4] for details) and $\mathscr{M}_{\mathbb{X}}$ stands for the corresponding Borel σ–algebra. A *completely random measure* (CRM) is a random element $\tilde{\mu}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $(\boldsymbol{M}_{\mathbb{X}}, \mathscr{M}_{\mathbb{X}})$ such that, for any k–tuple $A_1, \ldots, A_k$ of pairwise disjoint sets in $\mathscr{X}$ and for any $k \geq 2$, the random variables $\tilde{\mu}(A_1), \ldots, \tilde{\mu}(A_k)$ are mutually independent. If it is further assumed that $\tilde{\mu}$ does not have random masses at fixed locations, then

$$\tilde{\mu} = \sum_{i \geq 1} J_i \, \delta_{X_i}$$

where $\{(J_i, X_i) : \; i = 1, 2, \ldots\}$ are the points of a Poisson process N on $\mathbb{R}^+ \times \mathbb{X}$ with intensity measure ν, i.e. for any $A \in \mathscr{B}(\mathbb{R}^+) \otimes \mathscr{X}$ such that $\nu(A) < \infty$ one has $N(A) = \mathrm{card}(\{(J_i, X_i) : \; i = 1, 2, \ldots\} \cap A)$ and

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

where $\mathbb{1}_A$ is the indicator function of set A. Moreover, the measure ν is such that $\int_{\mathbb{R}^+ \times B} \min\{s, 1\} \, \nu(\mathrm{d}s, \mathrm{d}x) < \infty$, for any B in $\mathscr{X}$, and is also referred to as the *intensity*, or the *Lévy intensity*, of the CRM $\tilde{\mu}$. It is important to emphasize that it characterizes $\tilde{\mu}$ through its Laplace functional transform representation. Indeed, if $f : \mathbb{X} \to \mathbb{R}$ is any measurable function such that $\int |f| \, \mathrm{d}\tilde{\mu} < \infty$ almost surely, then

$$\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{X}} f(x)\,\tilde{\mu}(\mathrm{d}x)}\right] = \exp\left\{-\int_{\mathbb{R}^+\times\mathbb{X}} \left[1-\mathrm{e}^{-s\,f(x)}\right]\,\nu(\mathrm{d}s,\mathrm{d}x)\right\} \tag{11}$$

Given its relevance, in the sequel we shall refer to

$$\psi_\nu^*(f) = \int_{\mathbb{R}^+\times\mathbb{X}} \left[1-\mathrm{e}^{-s\,f(x)}\right]\,\nu(\mathrm{d}s,\mathrm{d}x) \tag{12}$$

as the *Laplace exponent* of $\tilde{\mu}$. A popular example is the so–called *gamma CRM* with base measure cP_0, where $c>0$ and P_0 is a probability measure on $(\mathbb{X},\mathscr{X})$. This is characterized by the Lévy intensity $\nu(\mathrm{d}s,\mathrm{d}x) = s^{-1}\,\mathrm{e}^{-s}\mathrm{d}s\,c\,P_0(\mathrm{d}x)$ which leads to the following form of the Laplace functional transform

$$\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{X}} f(x)\,\tilde{\mu}(\mathrm{d}x)}\right] = \exp\left\{-c\int_{\mathbb{X}} \log[1+f(x)]\,P_0(\mathrm{d}x)\right\} \tag{13}$$

for any measurable function $f:\mathbb{X}\to\mathbb{R}$ such that $\int\log(1+|f|)\,\mathrm{d}P_0<\infty$. If one sets $f=\lambda\mathbb{1}_A$ for any set A in $\mathscr{X}$ and $\lambda>0$, from (13) one deduces that $\tilde{\mu}(A)$ has Laplace transform as in (5) and is, thus, gamma distributed with shape and scale parameters equal to $cP_0(A)$ and 1, respectively. Another noteworthy example corresponds to $\nu(\mathrm{d}s,\mathrm{d}x) = \sigma s^{-1-\sigma}\,\mathrm{d}\sigma\,P_0(\mathrm{d}x)/\Gamma(1-\sigma)$, where $\sigma\in(0,1)$, and identifies the so–called *σ–stable* CRM. In this case, for any measurable function $f:\mathbb{X}\to\mathbb{R}^+$ such that $\int_{\mathbb{X}} f^\sigma\,\mathrm{d}P_0<\infty$ one has

$$\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{X}} f(x)\,\tilde{\mu}(\mathrm{d}x)}\right] = \exp\left\{-\int_{\mathbb{X}} f^\sigma(x)\,P_0(\mathrm{d}x)\right\}. \tag{14}$$

A choice of $f=\lambda\mathbb{1}_A$ shows that $\tilde{\mu}(A)$ has a positive stable distribution with parameter $\sigma\in(0,1)$.

As shown in [20], CRMs are the basic building blocks for the definition of nonparametric priors in an exchangeable framework. Indeed, suitable transformations of $\tilde{\mu}$ lead to define random probability measures whose distribution can be used as the mixing Q in (1). The transformation of interest for the present paper is the normalization. Indeed, if the intensity ν of $\tilde{\mu}$ is such that $\nu(\mathbb{R}^+\times\mathbb{X})=\infty$, it is possible to set Q as the probability distribution of

$$\tilde{p} = \frac{\tilde{\mu}}{\tilde{\mu}(\mathbb{X})}, \tag{15}$$

which defines a *normalized CRM* or a *normalized random measure with independent increments*. See [27] and [14]. For example, the Dirichlet process with base measure cP_0 can be obtained through (15) if $\tilde{\mu}$ is a gamma CRM whose Laplace functional transform is as in (13). Analogously, a shown in [16] the normalized σ–stable CRM with base measure P_0 coincides with the normalization of a CRM characterized through (14). The construction displayed in (15) is also relevant for our proposal. Indeed, we shall consider a vector $(\tilde{\mu}_1,\dots,\tilde{\mu}_n)$

of dependent CRMs such that marginally each $\tilde{\mu}_i$ is a gamma CRM. A vector of dependent random probabilities will, then, be obtained by normalizing $\tilde{\mu}_i$, for $i = 1, \dots, n$. The resulting dependent Dirichlet processes will represent a candidate for modelling partially exchangeable data as in (2) when the cardinality of $\mathcal{Z}$ is n.

3. Gamma CRM and Dirichlet process vectors

The definition outlined in Section 2 can be extended to the case of CRM vectors. These are random elements $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}_1, \dots, \tilde{\mu}_n)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $(\boldsymbol{M}_{\mathbb{X}}^n, \mathscr{M}_{\mathbb{X}}^n)$ such that, for any $k \geq 2$ and for any choice of pairwise disjoint sets $A_1, \dots, A_k$ in $\mathscr{X}$, the vectors $(\tilde{\mu}_1(A_i), \dots, \tilde{\mu}_n(A_i))$ are, for $i = 1, \dots, k$, mutually independent. In a similar fashion as for the one–dimensional case, a Poisson process type representation holds true and

$$(\tilde{\mu}_1, \dots, \tilde{\mu}_n) = \sum_r (J_{r,1}, \dots, J_{r,n})\, \delta_{X_r}.$$

In the above representation, $\{(J_{r,1}, \dots, J_{r,n}, X_r) : \ r = 1, 2, \dots\}$ are points from a Poisson process on $((\mathbb{R}^+)^n, \mathbb{X})$ with intensity measure ν such that

$$\int_{(\mathbb{R}^+)^n \times B} \|\boldsymbol{s}\|\, \nu(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) < \infty \qquad \forall B \in \mathscr{X}$$

where $\nu(A_{(i)} \times B) = \nu_i(A \times B)$ for any A in $\mathscr{B}(\mathbb{R}^+)$, ν_i is the (marginal) intensity of $\tilde{\mu}_i$, $A_{(i)} = (\mathbb{R}^+)^{i-1} \times A \times (\mathbb{R}^+)^{n-i}$ and $\|\boldsymbol{x}\|$ is the usual notation for the Euclidean norm of vector $\boldsymbol{x} \in \mathbb{R}^n$. Similarly to the one–dimensional case, ν characterizes $\tilde{\boldsymbol{\mu}}$ through its Laplace functional transform representation. Indeed, for any collection of measurable functions $f_i : \mathbb{X} \to \mathbb{R}$ such that $\int |f|\, \mathrm{d}\tilde{\mu}_i < \infty$, almost surely, $i = 1, \dots, n$, one has

$$\mathbb{E}\left[\mathrm{e}^{-\sum_{i=1}^n \int_{\mathbb{X}} f_i(x)\, \tilde{\mu}_i(\mathrm{d}x)}\right] = \exp\left\{-\int_{(\mathbb{R}^+)^n \times \mathbb{X}} \left[1 - \mathrm{e}^{\langle \boldsymbol{y}, \boldsymbol{f}(x)\rangle}\right]\, \nu(\mathrm{d}\boldsymbol{y}, \mathrm{d}x)\right\} \tag{16}$$

where $\boldsymbol{f} = (f_1, \dots, f_n)$, $\boldsymbol{y} = (y_1, \dots, y_n)$ and $\langle \boldsymbol{y}, \boldsymbol{f}(x)\rangle = \sum_{i=1}^n y_i f_i(x)$. In order to simplify the notation, we shall henceforth identify the Laplace exponent of $\tilde{\boldsymbol{\mu}}$ as

$$\psi^*_{\nu,n}(\boldsymbol{f}) = \int_{(\mathbb{R}^+)^n \times \mathbb{X}} \left[1 - \mathrm{e}^{\langle \boldsymbol{y}, \boldsymbol{f}(x)\rangle}\right]\, \nu(\mathrm{d}\boldsymbol{y}, \mathrm{d}x). \tag{17}$$

Example 1. Instances of multivariate CRMs with a view to applications in Bayesian nonparametric inference can be found, e.g., in [9] and in [17]. For example, a vector of CRMs $(\tilde{\mu}_1, \tilde{\mu}_2)$ with σ–stable margins having Laplace transform (14) has Lévy intensity of the form

$$\nu(\mathrm{d}y_1, \mathrm{d}y_2, \mathrm{d}x) = \frac{\sigma^2(1+\theta)}{\Gamma(1-\sigma)}\, \frac{y_1^{\theta\sigma-1} y_2^{\theta\sigma-1}}{\left\{y_1^{\theta\sigma} + y_2^{\theta\sigma}\right\}^{\frac{1}{\theta}+2}}\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}x$$

for any $\theta > 0$. The corresponding Laplace exponent coincides with

$$\frac{\psi^*_{\nu,2}(\lambda_1 \mathbb{1}_{A_1}, \lambda_2 \mathbb{1}_A)}{P_0(A)} = \lambda_1^\sigma + \lambda_2^\sigma - \frac{\lambda_1 \lambda_2}{\Gamma(1-\sigma)} \int_{(\mathbb{R}^+)^2} \frac{\mathrm{e}^{-\lambda_1 y_1 - \lambda_2 y_2}}{\left\{y_1^{\theta\sigma} + y_2^{\theta\sigma}\right\}^{\frac{1}{\theta}}} \, \mathrm{d}y_1 \, \mathrm{d}y_2$$

For further details, see [9] where this vector of dependent CRMs has been used to define neutral to the right priors for pairs of survival functions with two–sample data. An obvious simplification occurs when $\theta = 1/\sigma$. In this case, indeed, one has

$$\frac{\psi^*_{\nu,2}(\lambda_1 \mathbb{1}_{A_1}, \lambda_2 \mathbb{1}_A)}{P_0(A)} = \frac{\lambda_1^{\sigma+1} - \lambda_2^{\sigma+1}}{\lambda_1 - \lambda_2} \mathbb{1}_{\mathbb{R}\setminus\{0\}}(\lambda_1 - \lambda_2) + (\sigma+1)\lambda_1^\sigma \, \mathbb{1}_{\{0\}}(\lambda_1 - \lambda_2).$$

This simplifed structure has allowed the definition of a bivariate version of the two–parameter Poisson–Dirichlet process in [17]. □

In contrast to this example, we now wish to introduce a CRM vector $\tilde{\boldsymbol{\mu}}$ whose components marginally are gamma CRMs. This is accomplished by introducing the following multivariate Lévy intensity on $(\mathbb{R}^+)^n \times \mathbb{X}$

$$\nu(\mathrm{d}\boldsymbol{y}, \mathrm{d}x) = \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-i-1)!} \frac{e^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}} \, \mathrm{d}\boldsymbol{y} \, c \, P_0(\mathrm{d}x) \tag{18}$$

with $|\boldsymbol{y}| = \sum_{=1}^n y_i$. It is easy to check that

$$\nu(A_{(i)} \times B) = cP_0(B) \int_A \frac{\mathrm{e}^{-y}}{y} \, \mathrm{d}y \qquad i = 1, \ldots, n,$$

for any $A \in \mathscr{B}(\mathbb{R}^+)$, which implies that $\tilde{\mu}_1, \ldots, \tilde{\mu}_n$ marginally are identically distributed gamma CRMs. It is worth noting that (18) corresponds to the superposition of n vector of CRMs

$$(\tilde{\mu}_1, \ldots, \tilde{\mu}_n) = \sum_{i=1}^n (\mu^*_{i,1}, \ldots, \mu^*_{i,n})$$

with the intensity of the ith summand $(\mu^*_{i,1}, \ldots, \mu^*_{i,n})$ being

$$\nu_i^*(\mathrm{d}\boldsymbol{y}, \mathrm{d}x) = cP_0(\mathrm{d}x) \, \frac{(n-1)!}{(n-i)!} \, \frac{\mathrm{e}^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^i}$$

Moreover, one has $\nu_1^*((\mathbb{R}^+)^n \times B) = \infty$, whereas $\nu_i^*((\mathbb{R}^+)^n \times B) < \infty$ for any $i = 2, \ldots, n$ and for any $B \in \mathscr{X}$. Hence, for any $i \geq 2$ the vector $(\mu^*_{i,n}, \ldots, \mu^*_{n,n})$ has a finite number of jumps and acts as a multivariate Poisson compound process. This structure is reminiscent of a completely random measure whose normalization has been shown in [29] to be dense in the class of homogeneous normalized completely random measures.

The components of the gamma CRM vector will be normalized, thus yielding a random probability vector

$$(\tilde{p}_1, \dots, \tilde{p}_n) = \left(\frac{\tilde{\mu}_1}{\tilde{\mu}_1(\mathbb{X})}, \dots, \frac{\tilde{\mu}_n}{\tilde{\mu}_n(\mathbb{X})} \right) \tag{19}$$

that we shall refer to as n–variate Dirichlet process. Indeed, since each $\tilde{\mu}_i$ is a gamma process with base measure cP_0, then $\tilde{p}_i$ is a Dirichlet process with base measure cP_0, for any $i = 1, \dots, n$. Before discussing the uses of (19) for drawing inferences with partially exchangeable observations, we shall linger on some important distributional properties of the un-normalized gamma CRM vector $\tilde{\boldsymbol{\mu}}$.

4. Connection with Lévy copulas

A popular tool for the definition of multivariate distributions on $\mathbb{R}^d$, with fixed margins, is represented by copulas. See [25] for an extensive account on the topic. A similar approach has been recently introduced in an infinite-dimensional setting for the definition of multidimensional CRM vectors with fixed marginal CRMs. Lévy copulas have been originally employed for applications in mathematical finance, with the goal of describing the dynamics of portfolios including multiple dependent assets whose evolution in time is modeled through a Lévy process. See [3]. Dependence among the components of a Lévy process vector is, then, induced by means of Lévy copulas which act at the level of Lévy intensities. Since increasing Lévy processes are special case of CRMs, one can easily extend the use of Lévy copulas for defining CRM vectors, as illustrated in [9] and [17]. In view of the definition provided in the previous section, one might wonder whether it is possible to identify the Lévy copula that induces (18) starting from gamma univariate margins. Note that, according to Theorem 5.4 in [3] such a copula is unique. The interest in such a result is motivated by the fact that a possible representation of dependence through a Lévy copula allows one to simulate approximate realizations of the trajectories of the CRM vector. Indeed, this can be achieved by relying on an extension of the well–known Ferguson and Klass algorithm. For details, see Algorithm 6.5 in [3]. Besides this, it can provide further insight into the dependence structure featured by the CRM vector we have introduced.

Let us focus on the two–dimensional case, i.e. $n = 2$, and suppose each marginal CRM is homogeneous with intensity that can be represented as follows

$$\nu_i(\mathrm{d}s, \mathrm{d}x) = \rho_i(s)\, \mathrm{d}s\, cP_0(\mathrm{d}x). \tag{20}$$

Note that the our proposal clearly fits this setting. A Lévy copula is, then, a function $C : [0, \infty]^2 \to [0, \infty]$ such that

(i) $C(y_1, 0) = C(0, y_2) = 0$ for any positive y_1 and y_2,

(ii) C has uniform margins, i.e. $C(y_1, \infty) = y_1$ and $C(\infty, y_2) = y_2$,

(iii) for all $y_1 < z_1$ and $y_2 < z_2$, $C(y_1, y_2) + C(z_1, z_2) - C(y_1, z_2) - C(y_2, z_1) \geq 0$.

Set $y \mapsto U_i(y) := \int_y^\infty \rho_i(s)\,\mathrm{d}s$ as the i–th marginal tail integral associated to ρ_i. Moreover, if

$$\nu(\mathrm{d}s_1, \mathrm{d}s_2, \mathrm{d}x) = \rho(s_1, s_2)\,\mathrm{d}s_1\,\mathrm{d}s_2\, cP_0(\mathrm{d}x) \tag{21}$$

is the Lévy intensity of $(\tilde{\mu}_1, \tilde{\mu}_2)$ and $(y_1, y_2) \mapsto U(y_1, y_2) = \int_{y_1}^\infty \int_{y_2}^\infty \rho(s_1, s_2)\,\mathrm{d}s_1\,\mathrm{d}s_2$ is the corresponding tail integral. According to Theorem 5.4 in [3] there exists a unique Lévy copula C such that $U(y_1, y_2) = C(U_1(y_1), U_2(y_2))$. Furthermore, if both the copula C and the marginal tail integrals are sufficiently smooth

$$\rho(s_1, s_2) = \left.\frac{\partial^2 C(y_1, y_2)}{\partial y_1 \partial y_2}\right|_{y_1 = U_1(s_1),\, y_2 = U_2(s_2)} \rho_1(s_1)\,\rho_2(s_2).$$

A wide range of dependence structures can be induced through Lévy copulas. For example the independence case, i.e. $\int_{A\times B} \rho(s_1, s_2)\,\mathrm{d}s_1\,\mathrm{d}s_2 = \int_A \rho(s_1)\,\mathrm{d}s_1 + \int_B \rho_2(s_2)\,\mathrm{d}s_2$ for any A and B in $\mathscr{B}(\mathbb{R}^+)$, corresponds to the Lévy copula

$$C_\perp(y_1, y_2) = y_1 \mathbb{1}_{\{\infty\}}(y_2) + y_2 \mathbb{1}_{\{\infty\}}(y_1).$$

On the other hand, the case of completely dependent CRMs corresponds to

$$C_\parallel(y_1, y_2) = \min\{y_1, y_2\}$$

which yields a vector $(\tilde{\mu}_1, \tilde{\mu}_2)$ such that for any x and y in $\mathbb{X}$ either $\tilde{\mu}_i(\{x\}) < \tilde{\mu}_i(\{y\})$ or $\tilde{\mu}_i(\{x\}) > \tilde{\mu}_i(\{y\})$, for $i = 1, 2$, almost surely. Intermediate cases, between these two extremes, can be detected, for example, by relying on the *Lévy-Clayton* copula defined by

$$C_\theta(y_1, y_2) = (y_1^{-\theta} + y_2^{-\theta})^{-\frac{1}{\theta}} \qquad \theta > 0. \tag{22}$$

with the parameter θ regulating the degree of dependence. It can be seen that $\lim_{\theta\to 0} C_\theta = C_\perp$ and $\lim_{\theta\to\infty} C_\theta = C_\parallel$.

We shall now focus on the determination of the Lévy copula that identifies the specific gamma CRM vector characterized by the Lévy measure (18). We shall use the notation $\Gamma(a, x) = \int_x^\infty s^{a-1}\,\mathrm{e}^{-s}\,\mathrm{d}s$ for the incomplete gamma function, whereas $\Gamma^{-1}(a, x)$ is the inverse function of $x \mapsto \Gamma(a, x)$, for any $a \in \mathbb{R}$.

Proposition 1. *The measure ν defined in (18) with $n = 2$ can be recovered by applying the Lévy Copula*

$$C(y_1, y_2) = \Gamma(0, \Gamma^{-1}(0, y_1) + \Gamma^{-1}(0, y_2)) \tag{23}$$

to a pair of gamma CRMs.

Proof. The tail integral of each marginal gamma CRM is

$$\mathcal{U}_i(x) = \int_x^{+\infty} y^{-1} e^{-y}\,\mathrm{d}y = \Gamma(0, x) \qquad i = 1, 2$$

On the other hand, the tail integral associated to (18) with $n = 2$ is

$$\mathcal{U}(x_1, x_2) = \int_{x_1}^{+\infty} \int_{x_2}^{+\infty} \left[\frac{1}{(y_1 + y_2)^2} \mathrm{e}^{-y_1 - y_2} + \frac{1}{(y_1 + y_2)} \mathrm{e}^{-y_1 - y_2} \right] \mathrm{d}y_1 \, \mathrm{d}y_2$$

After the change of variable $s = y_1 + y_2$ and $t = y_1$ we obtain

$$\begin{aligned} \mathcal{U}(x_1, x_2) &= \int_{x_1+x_2}^{+\infty} e^{-s} \left(\frac{1}{s^2} + \frac{1}{s} \right) \int_{x_1}^{s-x_2} \mathrm{d}t \, \mathrm{d}s \\ &= \int_{x_1+x_2}^{+\infty} e^{-s} \left(\frac{s - (x_1 + x_2)}{s^2} + \frac{s - (x_1 + x_2)}{s} \right) \mathrm{d}s \\ &= \Gamma(0, x_1 + x_2) - (x_1 + x_2)\Gamma(-1, x_1 + x_2) + e^{-x_1 - x_2} \\ &\qquad - (x_1 + x_2)\Gamma(0, x_1 + x_2) \end{aligned}$$

Since $\Gamma(a + 1, x) = a\Gamma(a, x) + x^a \, \mathrm{e}^{-x}$ one has

$$\begin{aligned} \mathcal{U}(x_1, x_2) = \Gamma(0, x_1 + x_2) - (x_1 + x_2) &\left[\frac{e^{-x_1 - x_2}}{x_1 + x_2} - \Gamma(0, x_1 + x_2) \right] \\ &+ e^{-x_1 - x_2} - (x_1 + x_2)\Gamma(0, x_1 + x_2) \\ = \Gamma(0, x_1 + x_2) & \end{aligned}$$

From Theorem 5.3 in [3], the copula C for this process is characterized by

$$\mathcal{U}(x_1, x_2) = C(\mathcal{U}(x_1), \mathcal{U}(x_2))$$

which in this case reduces to

$$\Gamma(0, x_1 + x_2) = C(\Gamma(0, x_1), \Gamma(0, x_2))$$

Setting $y_i = \Gamma(0, x_i)$, for $i = 1, 2$, completes the proof. □

5. The Laplace exponent

As already mentioned, the Laplace exponent (12) of a CRM $\tilde{\mu}$ is an important tool for its possible uses in Bayesian nonparametric inference as pointed out, e.g., in [20]. For example, it may help to determine, via differentiation, a formula for the so-called *exchangeable partition probability function* (EPPF) corresponding to the distribution of a sample of exchangeable observations $(X_n)_{n \geq 1}$ as in (1) with Q being the probability distribution of a normalized CRM $\tilde{p} := \tilde{\mu}/\tilde{\mu}(\mathbb{X})$. The EPPF is also a key quantity needed for the posterior calculus of $\tilde{p}$, given the data. Similarly, the availability of an exact expression of ψ_ν^* is decisive for evaluating posterior inferences, either exact or approximated, with exchangeable survival data when $\mathbb{X} = \mathbb{R}^+$ and $\tilde{\mu}$ is both taken to define a neutral to the right prior for the survival function as

$$\mathbb{P}[X_i > t \,|\, \tilde{\mu}] = \tilde{S}(t) = \mathrm{e}^{-\tilde{\mu}(0,t]}$$

and as the mixing measure in a prior representation for the hazard rate function through

$$\tilde{S}(t) = \exp\left\{ - \int_0^t \int_{\mathbb{X}} k(s,x)\tilde{\mu}(\mathrm{d}x)\,\mathrm{d}s \right\}$$

for some kernel function $k(\,\cdot\,;\,\cdot\,)$ on $\mathbb{R}^+ \times \mathbb{X}$.

All these arguments are still relevant when working in a multivariate framework, where a multidimensional analogue of ψ^*_ν can be similarly applied to a Bayesian nonparametric model with partially exchangeable observations: this motivates the focus we are reserving in this Section to the determination of the Laplace exponent induced by (18).

Before getting started, it is worth noting that, since $(\tilde{\mu}_1, \ldots, \tilde{\mu}_n)$ has independent increments, its distribution is characterized by a choice of $f_1, \ldots, f_n$ in (16) such that $f_i = \lambda_i \, \mathbb{1}_A$ for any set A in $\mathscr{X}$, $\lambda_i \in \mathbb{R}^+$ and $i = 1, \ldots, n$. In this case

$$\psi^*_{\nu,n}(\boldsymbol{f}) = H(A)\,\psi_{\nu,n}(\boldsymbol{\lambda})$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \int_{(\mathbb{R}^+)^n} \left[1 - \mathrm{e}^{-\langle \boldsymbol{\lambda}, \boldsymbol{y}\rangle}\right] \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-i-1)!} \frac{e^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}}\,\mathrm{d}\boldsymbol{y} \tag{24}$$

Our goal is to prove the following

Proposition 2. *Let $\boldsymbol{\lambda} \in (\mathbb{R}^+)^n$ be such that it consists of $l \le n$ distinct values denoted as $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l$ with respective multiplicities $(n_1, \ldots, n_l)$. Then*

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \left(\prod_{i=1}^{l} \frac{1}{\Gamma(n_i)} \frac{\partial^{n_i - 1}}{\partial^{n_i - 1} \tilde{\lambda}_i} \right) \left(\phi_l(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l) \prod_{i=1}^{l} \tilde{\lambda}_i^{n_i - 1} \right), \tag{25}$$

where

$$\phi_n(\boldsymbol{x}) = \sum_{i=1}^{n} \frac{x_i^{n-1} \log(1 + x_i)}{\prod_{j=1, j\neq i}^{n} (x_i - x_j)} \mathbb{1}(x_1 \neq \ldots \neq x_n). \tag{26}$$

To prove Proposition 2, we first show in detail how to deal with the case $n = 2$ and, then, deduce the result for an arbitrary n by induction. A recursive formula provided in Lemma 1 will be the engine for the induction. For general n a delicate aspect we will encounter is related to deriving a form for $\psi_{\nu,n}(\boldsymbol{\lambda})$ when not all the coordinates in $\boldsymbol{\lambda}$ are distinct. For $n > 2$ for the sake of clarity it will be convenient to deal separately with a derivation when there are no ties in $\boldsymbol{\lambda}$ (Proposition 4) and then eventually to extend the proof to the case of general $\boldsymbol{\lambda} \in \mathbb{R}^n$. For the case $n = 2$, the joint Laplace functional transform is given as follows.

Proposition 3. *Let ν be the Lévy intensity introduced in (18) with $n = 2$. The corresponding Laplace exponent has the following form:*

$$\psi_{\nu,2}(\lambda_1, \lambda_2) = \begin{cases} [\lambda_1 \log(1 + \lambda_1) - \lambda_2 \log(1 + \lambda_2)]/(\lambda_1 - \lambda_2) & \lambda_1 \neq \lambda_2 \\ \log(1 + \lambda_1) + \lambda_1/(\lambda_1 + 1) & \lambda_1 = \lambda_2 \end{cases}$$

Proof. Suppose $\lambda_1 \neq \lambda_2$. Correspondingly one has

$$\psi_{\nu,2}(\lambda_1,\lambda_2) = I_1(\lambda_1,\lambda_2) + I_2(\lambda_1,\lambda_2)$$

where

$$I_1(\lambda_1,\lambda_2) = \int_0^\infty \int_0^\infty \left(1 - \mathrm{e}^{-\lambda_1 y_1 - \lambda_2 y_2}\right)\, dy_1\, dy_2\, \frac{\mathrm{e}^{-y_1-y_2}}{y_1+y_2}$$

$$I_2(\lambda_1,\lambda_2) = \int_0^\infty \int_0^\infty \left(1 - \mathrm{e}^{-\lambda_1 y_1 - \lambda_1 y_2}\right)\, dy_1\, dy_2\, \frac{\mathrm{e}^{-y_1-y_2}}{(y_1+y_2)^2}$$

The change of variable $y_1 + y_2 = w$ and $y_1/(y_1+y_2) = z$ leads to

$$\begin{aligned} I_1(s,t) &= \int_0^1 \mathrm{d}z \int_0^\infty \left(1 - e^{-w(\lambda_1 z + \lambda_2(1-z))}\right) \mathrm{e}^{-w}\, \mathrm{d}w \\ &= 1 - \frac{\log(1+\lambda_1) - \log(1+\lambda_2)}{\lambda_1 - \lambda_2} \end{aligned}$$

and, similarly

$$\begin{aligned} I_2(\lambda_1,\lambda_2) &= \int_0^1 \mathrm{d}z \int_0^\infty \left(1 - e^{-w(\lambda_1 z + \lambda_2(1-z))}\right) \frac{e^{-w}}{w}\, \mathrm{d}w \\ &= \frac{1+\lambda_1}{\lambda_1-\lambda_2} \log(1+\lambda_1) - \frac{1+\lambda_2}{\lambda_1-\lambda_2} \log(1+\lambda_2) - 1 \end{aligned}$$

and combining these two expression one obtains ψ_ν. Proceeding in a similar fashion, and with some useful simplifications, one also obtains $\psi_{\nu,2}(\lambda_1,\lambda_1)$ when $\lambda_1 = \lambda_2$. □

The statement of Proposition 2 points out that one needs to take into account possible ties in the vector $\boldsymbol{\lambda}$ when determining an expression, in closed form, of $\psi_{\nu,n}$. Hence, when dealing with the case $n > 2$ we shall first assume that $\boldsymbol{\lambda}$ has no ties and, then, move on to the case where any two λ_i and λ_j, with $i \neq j$, may coincide.

Set $E_n = \{\boldsymbol{x} \in (\mathbb{R}^+)^n : x_1 \neq x_2 \neq \cdots \neq x_n\}$. A preliminary Lemma provides a useful recursive formula for $\psi_{\nu,n}$ on E_n, with $n \geq 1$.

Lemma 1. *Suppose that $\boldsymbol{\lambda} \in E_{n+1}$, for any $n \geq 1$, and denote as $\boldsymbol{\lambda}_{-i}$ the original $\boldsymbol{\lambda}$ vector with the i–th component removed. Then the following recursive equation holds true*

$$\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \frac{\lambda_{n+1}}{\lambda_{n+1} - \lambda_n}\, \psi_{\nu,n}(\boldsymbol{\lambda}_{-n}) + \frac{\lambda_n}{\lambda_n - \lambda_{n+1}}\, \psi_{\nu,n}(\boldsymbol{\lambda}_{-(n+1)}) \tag{27}$$

Proof. If $A_j^n = \{\boldsymbol{k} \in \{0,1,\ldots,j\}^n : |\boldsymbol{k}| = j\}$, then

$$1 - \mathrm{e}^{-\langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle} = -\sum_{j \geq 1} \frac{(-1)^j (\langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle)^j}{j!}$$

$$= \sum_{j\ge 1} \frac{(-1)^{j+1}}{j!} \sum_{\boldsymbol{k}\in A_j^n} \frac{j!}{k_1!\,\cdots\,k_n!}\,\lambda_1^{k_1}\,\cdots\,\lambda_n^{k_n}\,y_1^{k_1}\,\cdots\,y_n^{k_n}$$

and

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!} \sum_{j\ge 1} \frac{(-1)^{j+1}}{j!} \sum_{\boldsymbol{k}\in A_j^n} \frac{j!}{k_1!\cdots k_n!} \lambda_1^{k_1}\cdots\lambda_n^{k_n}\, I_n^*(\boldsymbol{k})$$

where

$$I_n^*(\boldsymbol{k}) = \int_{(\mathbb{R}^+)^n} y_1^{k_1}\,\cdots\,y_n^{k_n}\,\frac{\mathrm{e}^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}}\,\mathrm{d}y_1\,\cdots\,\mathrm{d}y_n$$

A simple change of variable, namely $z_i = y_i/s$ for $i = 1,\dots,n-1$ and $s = |\boldsymbol{y}|$, yields

$$I_n^*(\boldsymbol{k}) = \frac{k_1!\cdots k_n!}{(j+n-1)!}(n-2-i+j)!.$$

This in turn leads to

$$\begin{aligned}
\psi_{\nu,n}(\boldsymbol{\lambda}) &= \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!} \sum_{j\ge 1} \sum_{\boldsymbol{k}\in A_j^n} \frac{(-1)^{j+1}\,(n-2-i+j)!}{(j+n-1)!}\,\lambda_1^{k_1}\,\cdots\,\lambda_n^{k_n} \\
&= (n-1)! \sum_{j\ge 1} \sum_{\boldsymbol{k}\in A_j^n} \frac{(-1)^{j+1}}{(j+n-1)!}\,\lambda_1^{k_1}\,\cdots\,\lambda_n^{k_n}\,\sum_{l=0}^{n-1} \frac{(l+j-1)!}{l!} \\
&= \sum_{j\ge 1} \sum_{\boldsymbol{k}\in A_j^n} \frac{(-1)^{j+1}}{j}\,\lambda_1^{k_1}\,\cdots\,\lambda_n^{k_n} \qquad (28)
\end{aligned}$$

since

$$\sum_{l=0}^{n-1} \frac{(l+j-1)!}{l!} = \frac{1}{j}\frac{(j+n-1)!}{(n-1)!}. \qquad (29)$$

Hence, if one resorts to (28)

$$\begin{aligned}
\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \sum_{j\ge 1} \frac{(-1)^{j+1}}{j} \sum_{k_1=0}^{j} \lambda_1^{k_1} \sum_{k_2=0}^{j-k_1} \lambda_2^{k_2}\,\cdots \\
\cdots \sum_{k_n=0}^{j-(k_1+\cdots+k_{n-1})} \lambda_n^{k_n}\,\lambda_{n+1}^{j-(k_1+\cdots+k_n)}
\end{aligned}$$

Afer some algebra, the last sum above can be rewritten as

$$\sum_{k_n=0}^{j-(k_1+\cdots+k_{n-1})} \lambda_n^{k_n}\,\lambda_{n+1}^{j-(k_1+\cdots+k_n)} = \lambda_{n+1}^{j-(k_1+\cdots+k_{n-1})} \sum_{k_n=0}^{j-(k_1+\cdots+k_{n-1})} \left(\frac{\lambda_n}{\lambda_{n+1}}\right)^{k_n}$$

$$= \frac{\lambda_{n+1}^{j-(k_1+\cdots+k_{n-1})+1} - \lambda_n^{j-(k_1+\cdots+k_{n-1})+1}}{\lambda_{n+1} - \lambda_n}$$

Hence

$$\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \frac{\lambda_{n+1}}{\lambda_{n+1} - \lambda_n} \sum_{j\geq 1} \frac{(-1)^{j+1}}{j} \sum_{k_1=0}^{j} \lambda_1^{k_1} \sum_{k_2=0}^{j-k_1} \lambda_2^{k_2} \cdots$$
$$\cdots \sum_{k_{n-1}=0}^{j-(k_1+\cdots+k_{n-2})} \lambda_{n-1}^{k_{n-1}} \lambda_{n+1}^{j-(k_1+\cdots+k_{n-1})}$$
$$+ \frac{\lambda_n}{\lambda_n - \lambda_{n+1}} \sum_{j\geq 1} \frac{(-1)^{j+1}}{j} \sum_{k_1=0}^{j} \lambda_1^{k_1} \sum_{k_2=0}^{j-k_1} \lambda_2^{k_2} \cdots$$
$$\cdots \sum_{k_{n-1}=0}^{j-(k_1+\cdots+k_{n-2})} \lambda_{n-1}^{k_{n-1}} \lambda_{n}^{j-(k_1+\cdots+k_{n-1})}$$

which shows the validity of (27). □

We are now in a position to state and prove the following representation of the Laplace exponent of our multivariate gamma CRM vector, when there are no ties in its argument $\boldsymbol{\lambda}$.

Proposition 4. *For any $\boldsymbol{\lambda} \in E_n$ and $n \geq 1$ one has*

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \phi_n(\boldsymbol{\lambda}) \tag{30}$$

where ϕ_n is as in (26).

Proof. Suppose (30) holds true for n and we shall show that this implies the validity of (30) for $n+1$. By virtue of Proposition 3 the proof is thus completed by induction. Since (30) holds true for n, for any $\boldsymbol{\lambda} \in (\mathbb{R}^+)^{n+1}$ one has

$$\psi_{\nu,n}(\boldsymbol{\lambda}_{-n}) = \frac{\lambda_{n+1}^{n-1}\log(1+\lambda_{n+1})}{\prod_{j=1}^{n-1}(\lambda_{n+1}-\lambda_j)} + \sum_{i=1}^{n-1} \frac{\lambda_i^{n-1}\log(1+\lambda_i)}{(\lambda_i - \lambda_{n+1})\prod_{j=1,j\neq i}^{n-1}(\lambda_i - \lambda_j)}$$

$$\psi_{\nu,n}(\boldsymbol{\lambda}_{-(n+1)}) = \sum_{i=1}^{n} \frac{\lambda_i^{n-1}\log(1+\lambda_i)}{\prod_{j=1,j\neq i}^{n}(\lambda_i - \lambda_j)}$$
$$= \frac{\lambda_n^{n-1}\log(1+\lambda_n)}{\prod_{j=1}^{n-1}(\lambda_n - \lambda_j)} + \sum_{i=1}^{n-1} \frac{\lambda_i^{n-1}\log(1+\lambda_i)}{(\lambda_i - \lambda_n)\prod_{j=1,j\neq i}^{n-1}(\lambda_i - \lambda_j)}$$

If these two expressions are plugged in the recursive relation (27) one has

$$\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \frac{\lambda_{n+1}^{n}\log(1+\lambda_{n+1})}{\prod_{j=1}^{n}(\lambda_{n+1}-\lambda_j)} + \frac{\lambda_n^{n}\log(1+\lambda_n)}{\prod_{j=1,j\neq n}^{n+1}(\lambda_n - \lambda_j)}$$

$$\sum_{i=1}^{n-1}\left[\frac{\lambda_{n+1}}{\lambda_i-\lambda_{n+1}}-\frac{\lambda_n}{\lambda_i-\lambda_n}\right]\frac{\lambda_i^{n-1}\log(1+\lambda_i)}{(\lambda_{n+1}-\lambda_n)\prod_{j=1,j\neq i}^{n-1}(\lambda_i-\lambda_j)}.$$

After some algebra, one shows that $\psi_{\nu,n+1}$ satisfies (30) and the proof is completed. $\square$

To achieve a complete proof of Proposition 2, it only remains to extend the technique used for the proof of Proposition 4, by taking this time into account possible ties in $\boldsymbol{\lambda}\in(\mathbb{R}^+)^n$.

Proof of Proposition 2. If $B_j=\{i:\ \lambda_i=\tilde\lambda_j\}$, for any $j=1,\dots,l$, and

$$|\boldsymbol{y}|_j=\sum_{i\in B_j}y_i$$

for any $\boldsymbol{y}\in(\mathbb{R}^+)^n$, one has, similarly to (28),

$$\psi_{\nu,n}(\boldsymbol{\lambda})=\sum_{i=0}^{n-1}\frac{(n-1)!}{(n-1-i)!}\sum_{j\geq 1}\frac{(-1)^{j+1}}{j!}\sum_{\boldsymbol{k}\in A_j^l}\frac{j!}{k_1!\cdots k_l!}\tilde\lambda_1^{k_1}\cdots\tilde\lambda_l^{k_l}\,I_n^{**}(\boldsymbol{k})$$

where

$$\begin{aligned}I_n^{**}(\boldsymbol{k})&=\int_{(\mathbb{R}^+)^n}|\boldsymbol{y}|_1^{k_1}\,\cdots\,|\boldsymbol{y}|_l^{k_l}\,\frac{\mathrm{e}^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}}\,\mathrm{d}y_1\,\cdots\,\mathrm{d}y_n\\&=\frac{(n-2-i+j)!}{(n+j-1)!}\,(n_1)_{k_1}\,\cdots\,(n_l)_{k_l}\end{aligned}$$

This implies that

$$\begin{aligned}\psi_{\nu,n}(\boldsymbol{\lambda})&=\sum_{i=0}^{n-1}\frac{(n-1)!}{(n-1-i)!}\sum_{j\geq 1}\frac{(-1)^{j+1}\,(n-2-i+j)!}{(n+j-1)!}\\&\qquad\times\sum_{\boldsymbol{k}\in A_j^l}\frac{(n_1)_{k_1}\cdots(n_l)_{k_l}}{k_1!\cdots k_l!}\tilde\lambda_1^{k_1}\,\cdots\,\tilde\lambda_l^{k_l}\\&=(n-1)!\sum_{j\geq 1}\frac{(-1)^{j+1}}{(n+j-1)!}\sum_{\boldsymbol{k}\in A_j^l}\frac{(n_1)_{k_1}\cdots(n_l)_{k_l}}{k_1!\cdots k_l!}\\&\qquad\times\ \tilde\lambda_1^{k_1}\cdots\tilde\lambda_l^{k_l}\sum_{l=0}^{n-1}\frac{(l+j-1)!}{l!}\\&=\sum_{j\geq 1}\frac{(-1)^{j+1}}{j}\sum_{\boldsymbol{k}\in A_j^l}\frac{(n_1)_{k_1}\cdots(n_l)_{k_l}}{k_1!\cdots k_l!}\,\tilde\lambda_1^{k_1}\,\cdots\,\tilde\lambda_l^{k_l}\end{aligned}$$

where the last equality follows from (29). Now note that

$$\frac{(n_i)_{k_i}}{k_i!}\lambda_i^{k_i} = \frac{(k_i+1)_{n_i-1}}{\Gamma(n_i)}\,\tilde{\lambda}_i^{k_i} = \frac{1}{\Gamma(n_i)}\frac{\partial^{n_i-1}}{\partial\tilde{\lambda}_i^{n_i-1}}\,\tilde{\lambda}_i^{n_i-1+k_i}$$

and from this deduce

$$\begin{aligned}
\psi_{\nu,n}(\boldsymbol{\lambda}) &= \sum_{j\ge 1}\frac{(-1)^{j+1}}{j}\sum_{\boldsymbol{k}\in A_j^l}\prod_{i=1}^{l}\frac{1}{\Gamma(n_i)}\,\frac{\partial^{n_i-1}}{\partial\tilde{\lambda}_i^{n_i-1}}\,\tilde{\lambda}_i^{n_i-1+k_i}\\
&= \frac{1}{\prod_{i=1}^{l}\Gamma(n_i)}\;\frac{\partial^{n-l}}{\partial\tilde{\lambda}_1^{n_1-1}\cdots\partial\tilde{\lambda}_l^{n_l-1}}\\
&\qquad\times\left(\tilde{\lambda}_1^{n_1-1}\cdots\tilde{\lambda}_l^{n_l-1}\sum_{j\ge 1}\frac{(-1)^{j+1}}{j}\sum_{\boldsymbol{k}\in A_j^l}\tilde{\lambda}_1^{k_1}\cdots\tilde{\lambda}_l^{k_l}\right)
\end{aligned}$$

which, from (28) and by virtue of the definition of the function ϕ_l in (30), completes the proof of (25). □

6. Investigating the dependence structure

The determination of the Laplace exponent discussed in the previous section is preliminary to an investigation of the dependence structure among the random probabilities of the vector $\tilde{\boldsymbol{p}} = (\tilde{p}_1, \ldots, \tilde{p}_n)$ of dependent Dirichlet processes defined by (19). Given $\tilde{\boldsymbol{p}}$ is an infinite–dimensional object, some hints on the dependence between any pair $\tilde{p}_i$ and $\tilde{p}_j$, for $i \ne j$, are conveyed by the correlation coefficient between $\tilde{p}_i(A)$ and $\tilde{p}_j(A)$, for any $A \in \mathscr{X}$. For this reason, we shall first provide an expression for such a coefficient and, then, rely on the same technique for determining the partially exchangeable partition probability function of the sample $\boldsymbol{X}^{q_1}(z_1), \ldots, \boldsymbol{X}^{q_n}(z_n)$ characterized through (2). Indeed, almost sure discreteness $\tilde{p}_i$ implies that $\mathbb{P}[X_i(z_\kappa) = X_j(z_\ell)] > 0$ for any i, j, κ and ℓ. In other terms, ties may appear both within each sample and between different samples $\boldsymbol{X}^{(q_\kappa)}(z_\kappa)$ and $\boldsymbol{X}^{(q_\ell)}(z_\ell)$. Hence, the $q_1 + \cdots + q_n$ data consist of k distinct values forming clusters of sizes $N_1, \ldots, N_k$ Moreover, $N_j = \sum_{i=1}^{n} q_{j,i} \ge 1$ with $q_{j,i}$ denoting the number of observations from $\boldsymbol{X}^{q_i}(z_i)$ coinciding with the j–th distinct value in the sample. In the following, we confine our treatment to the case $n = 2$.

6.1. Mixed moments and correlations

We rely on an approach used in [17] for defining a bivariate two–parameter Poisson–Dirichlet process: this arises as the normalization of a random measure that does not necessarily satisfy the property of independence when evaluated on pairwise disjoint sets as a CRM does. Accordingly, a key quantity is going to be

$$g_\rho(q_1, q_2, s, t) = \int_0^\infty \int_0^\infty y_1^{q_1} y_2^{q_2} e^{-sy_1 - ty_2} \rho(y_1, y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2 \tag{31}$$

where the notation corresponds to the one set forth in (20) and in (21): this is legitimate since the case we are examining is homogeneous. See (18). Note that one can also write

$$g_\rho(q_1, q_2; s, t) = I_1(q_1, q_2; s, t) + I_2(q_1, q_2; s, t)$$

where

$$I_1(q_1, q_2; s, t) = \int_{(\mathbb{R}^+)^2} y_1^{q_1}\, y_2^{q_2}\, \mathrm{e}^{-sy_1 - ty_2}\, \frac{\mathrm{e}^{-y_1 - y_2}}{(y_1 + y_2)^2}\, \mathrm{d}y_1\, \mathrm{d}y_2$$

$$I_2(q_1, q_2; s, t) = \int_{(\mathbb{R}^+)^2} y_1^{q_1}\, y_2^{q_2}\, \mathrm{e}^{-sy_1 - ty_2}\, \frac{\mathrm{e}^{-y_1 - y_2}}{y_1 + y_2}\, \mathrm{d}y_1\, \mathrm{d}y_2$$

when $q_1 + q_2 \geq 1$. Moreover, $g_\rho(0, 0; s, t) \equiv 1$. a simple change of variable into polar coordinates yields

$$\begin{aligned}
I_1(q_1, q_2, s, t) &= \int_0^{\frac{\pi}{2}} \sin(2\theta) \int_{\mathbb{R}^+} \rho^{q_1 + q_2 - 1} \cos^{2q_1}(\theta)\, \sin^{2q_2}(\theta) \\
&\qquad \times\ \mathrm{e}^{-\rho[(1+s)\cos^2(\theta) + (1+t)\sin^2(\theta)]} \mathrm{d}\rho\, \mathrm{d}\theta \\
&= \Gamma(q_1 + q_2) \int_0^{\frac{\pi}{2}} \frac{\cos^{2q_1}(\theta)\, \sin^{2q_2}(\theta)\, \sin(2\theta)}{[(1+s)\cos^2(\theta) + (1+t)\sin^2(\theta)]^{q_1 + q_2}}\, \mathrm{d}\theta \\
&= \Gamma(q_1 + q_2) \int_0^1 \frac{y^{q_1}(1-y)^{q_2}}{[(1+s)y + (1+t)(1-y)]^{q_1 + q_2}}\, \mathrm{d}y \\
&= (1+t)^{-q_1 - q_2} \Gamma(q_1 + q_2) \int_0^1 \frac{y^{q_1}(1-y)^{q_2}}{\left[1 - y\frac{t-s}{1+t}\right]^{q_1 + q_2}} \mathrm{d}y \\
&= \frac{\Gamma(q_2 + 1)\Gamma(q_1 + 1)}{(q_1 + q_2)(q_1 + q_2 + 1)}\, \frac{{}_2F_1(q_1 + q_2, q_1 + 1, q_1 + q_2 + 2, \frac{t-s}{1+t})}{(1+t)^{q_1 + q_2}}
\end{aligned}$$

In a similar fashion one determines $I_2(q_1, q_2; s, t)$ thus yielding

$$\begin{aligned}
g_\rho(q_1, q_2; s, t) &= \frac{\Gamma(q_2 + 1)\Gamma(q_1 + 1)}{(q_1 + q_2 + 1)(1+t)^{q_1 + q_2}} \\
&\qquad \times \left\{ \frac{{}_2F_1(q_1 + q_2, q_1 + 1, q_1 + q_2 + 2, \frac{t-s}{1+t})}{q_1 + q_2} \right. \\
&\qquad \left. + \frac{{}_2F_1(q_1 + q_2 + 1, q_1 + 1, q_1 + q_2 + 2, \frac{t-s}{1+t})}{1+t} \right\}
\end{aligned} \tag{32}$$

The availability of the g_ρ function allows us to determine an expression of the mixed moments of the un–normalized vector $(\tilde{\mu}_1(A), \tilde{\mu}_2(A))$, with $A \in \mathscr{X}$. This is a preliminary step towards the determination of the mixed moments of the normalized vector $(\tilde{p}_1(A), \tilde{p}_2(A))$. In the sequel, for any two vectors $\boldsymbol{x} = (x_1, \ldots, x_d)$ and $\boldsymbol{y} = (y_1, \ldots, y_d)$ in $\mathbb{N}_0^d$, one writes $\boldsymbol{x} \prec \boldsymbol{y}$ if either $|\boldsymbol{x}| < |\boldsymbol{y}|$ or $|\boldsymbol{x}| = |\boldsymbol{y}|$ and $x_1 < y_1$ or if $|\boldsymbol{x}| = |\boldsymbol{y}|$ with $x_i = y_i$ for $i = 1, \ldots, j$ and $x_{j+1} < y_{j+1}$ for some j in $\{1, \ldots, d\}$.

Proposition 5. *Let $p_j(q_1, q_2, k)$ be the set of vectors $(\boldsymbol{\lambda}, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_j)$ such that the coordinates of $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_j)$ are positive and such that $\sum_{i=1}^{j} \lambda_i = k$. Moreover, $\boldsymbol{s}_i = (s_{1,i}, s_{2,i})$ are vectors such that $\mathbf{0} \prec \boldsymbol{s}_1 \prec \cdots \prec \boldsymbol{s}_j$ and $\sum_{i=1}^{j} \lambda_i (s_{1,i} + s_{2,i}) = k = q_1 + q_2$. Then,*

$$\mathbb{E}\left[\prod_{i=1}^{2} \{\tilde{\mu}_i(A)\}^{q_i}\right] = q_1! q_2! \sum_{k=1}^{q_1+q_2} [H(A)]^k \times \sum_{j=1}^{q_1+q_2} \sum_{p_j(q_1,q_2,k)} \prod_{i=1}^{j} \frac{1}{\lambda_i! (s_{1,i} + s_{2,i})^{\lambda_i}}$$

Proof. Note that

$$\mathbb{E}\left[\mathrm{e}^{-s\tilde{\mu}_1(A) - t\tilde{\mu}_2(A)} \prod_{i=1}^{2} \{\tilde{\mu}_i(A)\}^{q_i}\right] = (-1)^{q_1+q_2} \frac{\partial^{q_1+q_2}}{\partial s^{q_1} \, \partial t^{q_2}} \, \mathrm{e}^{-H(A) \, \psi_{\rho,2}(s,t)}$$

and by virtue of Theorem 2.1 in [1] one has that the derivative in the right–hand side above coincides with

$$\begin{aligned} & \mathrm{e}^{-H(A)\psi_{\rho,2}(s,t)} \, q_1! q_2! \sum_{k=1}^{q_1+q_2} (-1)^k [H(A)]^k \times \\ & \qquad \times \sum_{j=1}^{q_1+q_2} \sum_{p_j(q_1,q_2,k)} \prod_{i=1}^{j} \frac{1}{\lambda_i! (s_{1,i}! s_{2,i}!)^{\lambda_i}} \left(\frac{\partial^{s_{1.i}+s_{2,i}}}{\partial s^{s_{1,i}} \partial t^{s_{2,i}}} \, \psi_{\rho,2}(s,t) \right)^{\lambda_i} \end{aligned}$$

By virtue of the definition of the function g_ν one has

$$\begin{aligned} & \mathrm{e}^{-H(A)\psi_{\rho,2}(s,t)} \, q_1! q_2! \sum_{k=1}^{q_1+q_2} [H(A)]^k \times \\ & \qquad \times \sum_{j=1}^{q_1+q_2} \sum_{p_j(q_1,q_2,k)} \prod_{i=1}^{j} \frac{1}{\lambda_i! (s_{1,i}! s_{2,i}!)^{\lambda_i}} \, \left(g_\rho(s_{1,i}, s_{2,i}, s, t) \right)^{\lambda_i}. \end{aligned}$$

Since $\psi_{\rho,2}(0,0) = 1$ and

$$g_\rho(s_{1,i}, s_{2,i}, 0, 0) = \frac{s_{1,i}! s_{2,i}!}{s_{1,i} + s_{2,i}}$$

the conclusion follows. ∎

The procedure described in the previous proof can be used to determine the correlation coefficient between $\tilde{p}_1(A)$ and $\tilde{p}_2(B)$, for any A and B in $\mathscr{X}$. In particular, when $A = B$, it will be seen that such a correlation does not depend on the specific set $\tilde{p}_1$ and $\tilde{p}_2$ are evaluated at: this finding is typically used to motivate the use of $\mathrm{corr}(\tilde{p}_1(A), \tilde{p}_2(A))$ as a measure of the dependence between $\tilde{p}_1$ and $\tilde{p}_2$. The main difference with respect to the previous proof is the use of a gamma integral representation of the normalizing total random masses $\tilde{\mu}_1(\mathbb{X})$ and $\tilde{\mu}_2(\mathbb{X})$. In order to make notation simpler, we set

$$
\begin{aligned}
h_1(z) &= {}_2F_1(2,2,4,z)\,{}_1F_1\left(c, c+1, -c(1-z)\frac{\log(1-z)}{z}\right) \\
h_2(z) &= {}_2F_1(3,2,4,z)\,{}_1F_1\left(c+1, c+2, -c(1-z)\frac{\log(1-z)}{z}\right)
\end{aligned}
\tag{33}
$$

where ${}_1F_1$ is the confluent hypergeometric function and ${}_2F_1$ is the Gauss hypergeometric function.

Proposition 6. *Let A and B be any two sets in $\mathscr{X}$ and suppose that $(\tilde{p}_1, \tilde{p}_2)$ is a Dirichlet vector defined as in (19), with $n = 2$. Then*

$$
\begin{aligned}
\mathrm{corr}(\tilde{p}_1(A), \tilde{p}_2(B)) = {}& \frac{[cH(A \cap B) - H(A)H(B)]}{3\sqrt{H(A)H(B)H(A^c)H(B^c)}} \\
& \times \Big\{ (c+1)\int_0^1 (1-z)^c \mathrm{e}^{-c\frac{z-1}{z}\log(1-z)}\, h_1(z)\,\mathrm{d}z \\
& + 2c\int_0^1 (1-z)^{c+1} \mathrm{e}^{-c\frac{z-1}{z}\log(1-z)}\, h_2(z)\,\mathrm{d}z \Big\}
\end{aligned}
\tag{34}
$$

where h_1 and h_2 are defined in (33).

Proof. Note, first, that

$$
\mathbb{E}\left[\tilde{p}_1(A)\tilde{p}_2(B)\right] = \int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde{\mu}_1(A)\tilde{\mu}_2(B)\,\mathrm{e}^{-s\tilde{\mu}_1(\mathbb{X}) - t\tilde{\mu}_2(\mathbb{X})}\right] \mathrm{d}u\,\mathrm{d}v. \tag{35}
$$

In order to exploit the independence of the increments of $(\tilde{\mu}_1, \tilde{\mu}_2)$ on can consider the (measurable) partition of $\mathbb{X}$ generated by $\{A, B\}$. Hence, one can rewrite the right–hand–side of (35) as follows

$$\begin{aligned}
&\int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde\mu_1(A\setminus B)\tilde\mu_2(B\setminus A)\,\mathrm{e}^{-s\tilde\mu_1(\mathbb{X})-t\tilde\mu_2(\mathbb{X})}\right]\,\mathrm{d}s\,\mathrm{d}t\\
&\quad+\int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde\mu_1(A\setminus B)\tilde\mu_2(A\cap B)\,\mathrm{e}^{-s\tilde\mu_1(\mathbb{X})-t\tilde\mu_2(\mathbb{X})}\right]\,\mathrm{d}s\,\mathrm{d}t\\
&\quad\quad+\int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde\mu_1(A\cap B)\tilde\mu_2(B\setminus A)\,\mathrm{e}^{-s\tilde\mu_1(\mathbb{X})-t\tilde\mu_2(\mathbb{X})}\right]\,\mathrm{d}s\,\mathrm{d}t\\
&\quad\quad\quad+\int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde\mu_1(A\cap B)\tilde\mu_2(A\cap B)\,\mathrm{e}^{-s\tilde\mu_1(\mathbb{X})-t\tilde\mu_2(\mathbb{X})}\right]\,\mathrm{d}s\,\mathrm{d}t
\end{aligned}\tag{36}$$

Let us focus on the first summand and note that it can be rewritten as

$$\begin{aligned}
\int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde\mu_1(A\setminus B)\,\mathrm{e}^{-s\tilde\mu_1(A\setminus B)-t\tilde\mu_2(A\setminus B)}\right] E\left[\tilde\mu_2(B\setminus A)\,\mathrm{e}^{-s\tilde\mu_1(B\setminus A)-t\tilde\mu_2(B\setminus A)}\right]\\
\times\ \mathbb{E}\left[\mathrm{e}^{-s\tilde\mu_1(\mathbb{X}^*)-t\tilde\mu_2(\mathbb{X}^*)}\right]\,\mathrm{d}s\,\mathrm{d}t
\end{aligned}$$

where $\mathbb{X}^*=\mathbb{X}\setminus(A\Delta B)$ and Δ stands for the symmetric difference between sets. Each factor in the integrand can be easily evaluated. As for the first one, for example, one has

$$\mathbb{E}\left[\tilde\mu_1(A\setminus B)\,\mathrm{e}^{-s\tilde\mu_1(A\setminus B)-t\tilde\mu_2(A\setminus B)}\right]=H(A\setminus B)\,\mathrm{e}^{-H(A\setminus B)\,\psi_{\rho,2}(s,t)}\,g_\rho(1,0;s,t)$$

and one, then, has

$$\begin{aligned}
&\int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde\mu_1(A\setminus B)\tilde\mu_2(B\setminus A)\,\mathrm{e}^{-s\tilde\mu_1(\mathbb{X})-t\tilde\mu_2(\mathbb{X})}\right]\,\mathrm{d}s\,\mathrm{d}t\\
&\qquad=H(A\setminus B)\,H(B\setminus A)\int_{(\mathbb{R}^+)^2}\mathrm{e}^{-c\psi_\rho(s,t)}\,g_\rho(1,0;s,t)\,g_\rho(0,1;s,t)\,\mathrm{d}s\,\mathrm{d}t
\end{aligned}$$

As for the last summand in (36), it should be noted that

$$\begin{aligned}
&\int_{(\mathbb{R}^+)^2} \mathbb{E}\left[\tilde\mu_1(A\cap B)\tilde\mu_2(A\cap B)\,\mathrm{e}^{-s\tilde\mu_1(\mathbb{X})-t\tilde\mu_2(\mathbb{X})}\right]\,\mathrm{d}s\,\mathrm{d}t\\
&\qquad=H(A\cap B)\int_{(\mathbb{R}^+)^2} g_\rho(1,1;s,t)\,\mathrm{e}^{-c\psi_\rho(s,t)}\,\mathrm{d}s\,\mathrm{d}t\\
&\qquad\quad+H^2(A\cap B)\int_{(\mathbb{R}^+)^2} g_\rho(1,0;s,t)\,g_\rho(0,1;s,t)\,\mathrm{e}^{-c\psi_\rho(s,t)}\,\mathrm{d}s\,\mathrm{d}t
\end{aligned}$$

Combining (35) and (36) with the above integral representations, one obtains

$$\begin{aligned}
&\mathbb{E}\left[\tilde{p}_1(A)\,\tilde{p}_2(B)\right]\\
&\quad = H(A)H(B)\int_{(\mathbb{R}^+)^2} g_\rho(1,0;s,t)\,g_\rho(0,1;s,t)\,\mathrm{e}^{-c\psi_\rho(s,t)}\,\mathrm{d}s\,\mathrm{d}t\\
&\qquad + H(A\cap B)\int_{(\mathbb{R}^+)^2} g_\rho(1,1;s,t)\,\mathrm{e}^{-c\psi_\rho(s,t)}\,\mathrm{d}s\,\mathrm{d}t
\end{aligned} \tag{37}$$

If $A = B = \mathbb{X}$ in (37), the following useful identity holds true

$$\begin{aligned}
&\int_{(\mathbb{R}^+)^2} g_\rho(1,0;s,t)\,g_\rho(0,1;s,t)\,\mathrm{e}^{-c\psi_\rho(s,t)}\,\mathrm{d}s\,\mathrm{d}t\\
&\qquad\qquad = \frac{1}{c^2} - \frac{1}{c}\int_{(\mathbb{R}^+)^2} g_\rho(1,1;s,t)\,\mathrm{e}^{-c\psi_\rho(s,t)}\,\mathrm{d}s\,\mathrm{d}t
\end{aligned}$$

which, in turn, yields

$$\begin{aligned}
\mathbb{E}\left[\tilde{p}_1(A)\,\tilde{p}_2(B)\right] &= \frac{H(A)H(B)}{c^2} + \left\{H(A\cap B) - \frac{H(A)H(B)}{c}\right\}\\
&\quad\times \int_{(\mathbb{R}^+)^2} g_\rho(1,1;s,t)\,\mathrm{e}^{-c\psi_\rho(s,t)}\,\mathrm{d}s\,\mathrm{d}t
\end{aligned} \tag{38}$$

Note, now, that from (32) one has

$$\int_{(\mathbb{R}^+)^2} \mathrm{e}^{-c\psi(s,t)} g_\nu(1,1,s,t)dsdt = J_1 + J_2$$

where

$$\begin{aligned}
J_1 &= \frac{1}{6}\int_{(\mathbb{R}^+)^2} \mathrm{e}^{-c\psi(s,t)}(1+t)^{-2}\,{}_2F_1\left(2,2,4,\frac{t-s}{1+t}\right)\mathrm{d}s\,\mathrm{d}t\\
J_2 &= \frac{1}{3}\int_{(\mathbb{R}^+)^2} \mathrm{e}^{-c\psi(s,t)}(1+t)^{-3}\,{}_2F_1\left(3,2,4,\frac{t-s}{1+t}\right)\mathrm{d}s\,\mathrm{d}t
\end{aligned}$$

A simple change of variable yields

$$\begin{aligned}
J_1 &= \frac{2}{6}\int_0^1\int_0^{1-z} w^{c-1}\mathrm{e}^{-cw\frac{\log(1-z)}{z}}\mathrm{e}^{-c\frac{z-1}{z}\log(1-z)}\,{}_2F_1(2,2,4,z)\,\mathrm{d}w\,\mathrm{d}z\\
&= \frac{1}{3}\int_0^1\left(\int_0^1 y^{c-1}\mathrm{e}^{-c(1-z)y\frac{\log(1-z)}{z}}\,\mathrm{d}y\right)(1-z)^c\mathrm{e}^{-c\frac{z-1}{z}\log(1-z)}\,{}_2F_1(2,2,4,z)\,\mathrm{d}z\\
&= \frac{B(1,c)}{3}\int_0^1 (1-z)^c\mathrm{e}^{-c\frac{z-1}{z}\log(1-z)}\,h_1(z)\,\mathrm{d}z
\end{aligned}$$

where, for any $a, b > 0$, $B(a, b)$ is the beta function. In a similar fashion one determines J_2. Finally, if one recalls that $\tilde{p}_1(A)$ and $\tilde{p}_2(B)$ are, marginally, beta distributed with parameters $(H(A), c - H(A))$ and $(H(B), c - H(B))$, the result follows. □

From (34) it follows that when $A = B$, the dependence on the set A disappears, as anticipated in (10), and $\mathrm{corr}(\tilde{p}_1(A), \tilde{p}_2(A))$ can be meant as a measure of the dependence between $\tilde{p}_1$ and $\tilde{p}_2$. This is a typical argument used in the Bayesian nonparametric literature.

6.2. Partially exchangeable random partition

As already mentioned, the components of the vector $(\tilde{p}_1, \tilde{p}_2)$ are such that ties may be detected both within and between the two samples $\boldsymbol{X}^{(q_1)}(z_1)$ and $\boldsymbol{X}^{(q_2)}(z_2)$. Hence, the sample induces a random partition of the integers $\{1, \ldots, q_1 + q_2\}$ that can be described through the so–called partially exchangeable partition probability function. This is the bivariate counterpart of the marginal exchangeable partition probability functions induced by $\tilde{p}_1$ and $\tilde{p}_2$. Indeed, since $\tilde{p}_1 \overset{d}{=} \tilde{p}_2$ and are Dirichlet processes with base measure $H = cP_0$, each sample $\boldsymbol{X}^{(q_i)}(z_i)$ characterized through (1) induces a a random partition that can be characterized through the probability function

$$\Pi_k^{(q_i)}(n_1, \ldots, n_k) = \frac{c^k}{(c)_{q_i}} \prod_{j=1}^{k} (n_j - 1)!$$

where $k \in \{1, \ldots, q_i\}$ is the number of sets of the partition and the positive integers $n_1, \ldots, n_k$ are the cardinalities of the partitions sets, thus being such that $\sum_{j=1}^{k} n_j = q_i$. The above displayed equation corresponds to the EPPF of the Dirichlet process whose base measure has total mass $c > 0$. The EPPF characterizing discrete random probabilities, even beyond the Dirichlet case, are a remarkable tool for addressing a variety of issues in Bayesian nonparametric inference. Indeed, they are the key for devising Blackwell-MacQueen type algorithms for density estimation (see [20] for a general discussion) and for the exact evaluation of estimators of interest in species sampling problems (see, e.g., [18, 10, 11]).

Here we deal an extension to the case where one considers the partition jointly induced by the two samples $\boldsymbol{X}^{(q_1)}(z_1)$ and $\boldsymbol{X}^{(q_2)}(z_2)$. The dependence between $\tilde{p}_1$ and $\tilde{p}_2$ obviously introduces some further technical issues with respect to the marginal exchangeable case. In order to examine such an extension in some detail, for any vectors $\boldsymbol{n}_1 = (n_{1,1}, \ldots, n_{k,1})$ and $\boldsymbol{n}_2 = (n_{1,2}, \ldots, n_{k,2})$ of non–negative integers in the set

$$\Delta_k(q_1, q_2) := \{(\boldsymbol{n}_1, \boldsymbol{n}_2) \in \mathbb{N}_0^{2k} : \; n_{j,1} + n_{j,2} \geq 1, \; |\boldsymbol{n}_i| = q_i\}$$

we shall denote by

$$\Pi_k^{(n_1,n_2)}(\boldsymbol{n}_1,\boldsymbol{n}_2) = \int_{\mathbb{X}^k} \mathbb{E}\left[\prod_{j=1}^k \tilde{p}_1^{n_{j,1}}(\mathrm{d}x_i)\,\tilde{p}_2^{n_{j,2}}(\mathrm{d}x_i)\right]$$

the probability of detecting two specific samples $\boldsymbol{X}^{n_1}(z_1)$ and $\boldsymbol{X}^{n_2}(z_2)$ featuring k distinct values with respective frequencies $n_{1,1}+n_{1,2},\dots,n_{k,1}+n_{k,2}$. Before providing an expression for $\Pi_k^{(n_1,n_2)}$, we need to introduce some notation. In particular, we set

$$\theta_i(q_1,q_2;z) = \begin{cases} \frac{q_1!q_2!}{(q_1+q_2+1)(q_1+q_2)}\,{}_2F_1(q_1+q_2,q_1+1;q_1+q_2+2;z) & i=1 \\ \frac{q_1!q_2!}{(q_1+q_2+1)}\,{}_2F_1(q_1+q_2+1,q_1+1;q_1+q_2+2;z) & i=0 \end{cases} \tag{39}$$

for $i=0,1$ and for any $z\in(0,1)$. We are now in a position to provide an expression of the pEPPF characterizing a bivariate Dirichlet process.

Proposition 7. *For any positive integers q_1, q_2 and k such that $k\le n_1+n_2$ and for any vector of cluster frequencies $(\boldsymbol{n}_1,\boldsymbol{n}_2)\in\Delta_k(n_1,n_2)$ one has*

$$\begin{aligned}
&\Pi_k^{(n_1,n_2)}(\boldsymbol{n}_1,\boldsymbol{n}_2)\\
&= \frac{c^k}{\prod_{i=1}^2\Gamma(n_i)} \sum_{\boldsymbol{i}\in\{0,1\}^k}\sum_{\ell=0}^{n_1-1}\sum_{m=0}^{n_2-1}\binom{n_1-1}{\ell}\binom{n_2-1}{m}(-1)^{k-|\boldsymbol{i}|+1}\\
&\quad\times\int_0^1 (1-z)^{c+n_1+m+k-|\boldsymbol{i}|-1-c\frac{z-1}{z}}\\
&\qquad\times\frac{{}_1F_1(\zeta(\ell,m,\boldsymbol{i}),\zeta(\ell,m,\boldsymbol{i})+1,-c(1-z)\frac{\log(1-z)}{z})}{\zeta(\ell,m,\boldsymbol{i})}\\
&\qquad\times\left(\prod_{j=1}^k\theta_{i_j}(n_{j,1},n_{j,2};z)+\prod_{j=1}^k\theta_{i_j}(n_{j,2},n_{j,1};z)\right)\mathrm{d}z
\end{aligned}\tag{40}$$

where $\zeta(\ell,m,\boldsymbol{i})=c+\ell+m+k-|\boldsymbol{i}|$ and $|\boldsymbol{i}|=i_1+\cdots+i_k$.

Proof. The result can be deduced from

$$\begin{aligned}\Pi_k^{(n_1,n_2)}(\boldsymbol{n}_1,\boldsymbol{n}_2) = \frac{c^k}{\prod_{i=1}^2\Gamma(n_i)}\left(\int_{A_-}+\int_{A_+}\right) s^{n_1-1}\,t^{n_2-1}\,\mathrm{e}^{-c\psi(s,t)}\\
\times\prod_{j=1}^k g_\nu(n_{j,1},n_{j,2};s,t)\,\mathrm{d}s\,\mathrm{d}t =: I_1+I_2\end{aligned}$$

where the function g_ν is as in (32), $A_- := \{(s,t)\in(\mathbb{R}^+)^2 : s<t\}$ and $A_+ := \{(s,t)\in(\mathbb{R}^+)^2 : s\ge t\}$. We shall explicitly deal with I_1, which is

associated to A_-, since an expression for I_2 can be similarly obtained. Resort to the change of variable $z = (t-s)/(1+t)$ and $w = 1/(1+t)$ and note that (z,w) is in the simplex $\mathcal{S}_1 = \{(z,w) \in [0,1]^2 : z+w \leq 1\}$ since $(s,t) \in A_-$. Hence

$$
\begin{aligned}
I_1 = {} & \frac{c^k}{\prod_{i=1}^2 \Gamma(n_i)} \sum_{|\boldsymbol{i}| \in \{0,1\}^k} \int_{\mathcal{S}_1} w^{c+k-|\boldsymbol{i}|-1} (1-z-w)^{n_1-1} (1-w)^{n_2-1} \\
& \times \mathrm{e}^{-c\frac{z-1}{z}\log(1-z) - cw\frac{\log(1-z)}{z}} \prod_{j=1}^k \theta_{i_j}(n_{j,1}, n_{j,2}, z)\, \mathrm{d}z\, \mathrm{d}w \\
= {} & \frac{c^k}{\prod_{i=1}^2 \Gamma(n_i)} \sum_{|\boldsymbol{i}| \in \{0,1\}^k} \sum_{\ell=0}^{n_1-1} \sum_{m=0}^{n_2-1} \binom{n_1-1}{\ell} \binom{n_2-1}{m} (-1)^{\ell+m} \\
& \times \int_{\mathcal{S}_1} w^{\zeta(\ell,m,\boldsymbol{i})-1} (1-z)^{n_1-1-\ell} \mathrm{e}^{-c\frac{z-1}{z}\log(1-z) - cw\frac{\log(1-z)}{z}} \\
& \times \prod_{j=1}^k \theta_{i_j}(n_{j,1}, n_{j,2}, z)\, \mathrm{d}z\, \mathrm{d}w
\end{aligned}
$$

and the first part in the representation of $\Pi_k^{(n_1,n_2)}$ follows upon noting that

$$
\begin{aligned}
& \int_0^{1-z} w^{\zeta(\ell,m,\boldsymbol{i})-1}\, \mathrm{e}^{-cw\frac{\log(1-z)}{z}}\, \mathrm{d}w \\
& \quad = (1-z)^{\zeta(\ell,m,\boldsymbol{i})} \int_0^1 y^{\zeta(\ell,m,\boldsymbol{i})-1}\, \mathrm{e}^{-cy(1-z)\frac{\log(1-z)}{z}}\, \mathrm{d}y \\
& \quad = (1-z)^{\zeta(\ell,m,\boldsymbol{i})} \frac{{}_1F_1(\zeta(\ell,m,\boldsymbol{i}), \zeta(\ell,m,\boldsymbol{i})+1, -c(1-z)\frac{\log(1-z)}{z})}{\zeta(\ell,m,\boldsymbol{i})}
\end{aligned}
$$

A similar procedure on A_+ leads to an expression of I_2 that completes the proof. $\square$

Acknowledgements

The authors are very grateful to an Associate Editor and two referees for their careful reading and valuable comments that have led to an improvement of the presentation. Antonio Lijoi's research is supported by the European Research Council (ERC) through StG "N-BNP" 306406. Fabrizio Leisen's research is partially supported by grant ECO2011-25706 of the Spanish Ministry of Science and Innovation. Dario Spanò's research is supported in part by CRiSM, an EPSRC-HEFCE UK grant.

References

[1] Constantines, G.M. and Savits, T.H. (1996). A multivariate version of the Faa di Bruno formula. *Trans. Amer. Math. Soc.* **348**, 503–520. MR1325915

[2] Cifarelli, D.M. and Regazzini, E. (1978). Problemi statistici non parametrici in condizioni di scambiabilità parziale. *Quaderni Istituto Matematica Finanziaria*, Università di Torino Serie III. English translation available at: http://www.unibocconi.it/wps/allegatiCTP/CR-Scamb-parz[1].20080528.135739.pdf

[3] Cont, R. and Tankov, P. (2004). *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL. MR2042661

[4] Daley, D.J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. 1.* Springer, New York. MR0950166

[5] de Finetti, B. (1938). Sur la condition de "équivalence partielle". In *Actualités Scientifique et Industrielle* **739**, 5–18. Hermann, Paris.

[6] De Iorio, M., Müller, P., Rosner, G.L. and MacEachern, S.N. (2004). An ANOVA model for dependent random measures. *J. Amer. Statist. Assoc.* **99**, 205–215. MR2054299

[7] Dunson, D.B., Xue, Y. and Carin, L. (2008). The matrix stick-breaking process: flexible Bayes meta-analysis. *J. Amer. Statist. Assoc.* **103**, 317–327. MR2420236

[8] Dunson, D.B. and Park, J.-H. (2008). Kernel stick-breaking processes *Biometrika* **95**, 307–323. MR2521586

[9] Epifani, I. and Lijoi, A. (2010). Nonparametric priors for vectors of survival functions. *Statistica Sinica* **20**, 1455–1484. MR2777332

[10] Favaro, S., Lijoi, A. and Prünster, I. (2012). Conditional formulae for Gibbs–type exchangeable random partitions. *Ann. Appl. Probab.*, to appear.

[11] Favaro, S., Lijoi, A. and Prünster, I. (2012). A new estimator of the discovery probability. *Biometrics* **68**, 1188–1196.

[12] Griffin, J.E. and Steel, M.F.J. (2006). Order-based dependent Dirichlet processes. *J. Amer. Statist. Assoc.* **101**, 179–194. MR2268037

[13] Hjort, N.L., Holmes, C.C., Müller, P. and Walker, S.G. (Eds.) (2010). *Bayesian Nonparametrics.* Cambridge University Press. MR2722987

[14] 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

[15] Jara, A., Hanson, T., Quintana, F.A., Müller, P. and Rosner, G. (2011). DPpackage: Bayesian non- and semi-parametric modelling in R. *J. Stat. Soft.* **40**, 1–30.

[16] Kingman, J.F.C. (1975). Random discrete distributions (with discussion). *J. Roy. Statist. Soc. Ser. B* **37**, 1–22. MR0368264

[17] Leisen, F. and Lijoi, A. (2011). Vectors of Poisson-Dirichlet processes. *J. Multivariate Anal.* **102**, 482–495. MR2755010

[18] Lijoi, A., Mena, R.H. and Prünster I. (2007). Bayesian nonparametric estimation of the probability of discovering new species. *Biometrika* **94**, 769–786. MR2416792

[19] Lijoi, A., Nipoti, B. and Prünster, I. (2011). Bayesian inference with dependent normalized completely random measures. *Carlo Alberto Notebooks* **224**.

[20] Lijoi, A. and Prünster, I. (2010). Beyond the Dirichlet process. *Bayesian Nonparametrics* (Holmes, C.C., Hjort, N.L., Müller, P. and Walker, S.G., Eds.), 80–136, Cambridge University Press, Cambridge. MR2730661

[21] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12**, 351–357. MR0733519

[22] MacEachern, S.N. (1999). Dependent nonparametric processes. In *ASA Proceedings of the Section on Bayesian Statistical Science*, Alexandria, VA: American Statistical Association.

[23] MacEachern, S.N. (2000). Dependent Dirichlet processes. *Technical Report*, Ohio State University.

[24] Müller, P., Quintana, F. and Rosner, G. (2004). A method for combining inference across related nonparametric Bayesian models. *J. R. Stat. Soc. Ser. B* **66**, 735–749. MR2088779

[25] Nelsen, R.B. (2006). *An introduction to copulas*. Springer, New York. MR2197664

[26] Rao, V.A. and Teh, Y.W. (2009). Spatial normalized Gamma processes. In *Advances in Neural Information Processing Systems* **22**.

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

[28] Rodríguez, A., Dunson, D. and Gelfand, A. (2008). The nested Dirichlet process. *J. Amer. Statist. Assoc.* **103**, 1131–1144. MR2528831

[29] Trippa, L. and Favaro, S. (2012). A class of normalized random measures with an exact predictive sampling scheme. *Scand. J. Statist.* **39**, 440–460.