

A conjugate class of random probability measures based on tilting and with its posterior analysis

JOHN W. LAU

School of Mathematics and Statistics, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia.

E-mail: john.lau@uwa.edu.au

This article constructs a class of random probability measures based on exponentially and polynomially tilting operated on the laws of completely random measures. The class is proved to be conjugate in that it covers both prior and posterior random probability measures in the Bayesian sense. Moreover, the class includes some common and widely used random probability measures, the normalized completely random measures (James (Poisson process partition calculus with applications to exchangeable models and Bayesian nonparametrics (2002) Preprint), Regazzini, Lijoi and Prünster (*Ann. Statist.* **31** (2003) 560–585), Lijoi, Mena and Prünster (*J. Amer. Statist. Assoc.* **100** (2005) 1278–1291)) and the Poisson–Dirichlet process (Pitman and Yor (*Ann. Probab.* **25** (1997) 855–900), Ishwaran and James (*J. Amer. Statist. Assoc.* **96** (2001) 161–173), Pitman (In *Science and Statistics: A Festschrift for Terry Speed* (2003) 1–34 IMS)), in a single construction. We describe an augmented version of the Blackwell–MacQueen Pólya urn sampling scheme (Blackwell and MacQueen (*Ann. Statist.* **1** (1973) 353–355)) that simplifies implementation and provide a simulation study for approximating the probabilities of partition sizes.

Keywords: Bayesian non-parametric; completely random measures; Dirichlet process; generalized gamma process; Poisson Dirichlet process; random probability measures; tilting

1. Introduction

Random probability measures derived from normalized independent increment processes have been studied for decades. Kingman [27] considers normalization over subordinators of Lévy processes with only positive jumps. Regazzini *et al.* [44] introduce the class of the normalized random measures with independent increments (normalized *completely random measures*) for studying the probabilistic properties of mean functionals of random probability measures. Investigations for statistical modeling are available in James [18], Lijoi *et al.* [32] and James *et al.* [22].

In Bayesian non-parametric statistics, the normalized random process is considered to be an unknown parameter and the posterior distribution of the process is usually of interest. The most popular class of such random processes for statistical modelling is the Dirichlet process (Ferguson [12]; Lo [36]). The Dirichlet process is appealing because it induces model flexibility and it is also conjugate in the sense that the posterior process, the process conditional on the data, is also a Dirichlet process. In fact, a surprising result shown in James *et al.* [21] states that only the Dirichlet process has the conjugacy property among the random probability measures in the class

of normalized independent increment processes. In the present work, we are able to show that it is not the case for a richer class of normalized processes. The class of normalized processes considered in this article is derived based on tilting, in particular, exponentially and polynomially tilting operated on the laws of completely random measures and this yields the class of laws containing the prior random probability measures and their posteriors in the Bayesian sense. So, the random probabilities in this class is conjugate.

Tilting the laws of random processes provides a way to enrich the class of random processes through change of measure. For example, Pitman [41] constructs the Poisson–Kingman process by normalizing a random process that has a tilted law of completely random measures. Other well known special cases are the Poisson–Dirichlet process, whose law is constructed by polynomially tilting the laws of positive α -stable processes (Pitman [41]), and the beta-gamma process (James [20]), whose law is obtained by polynomially tilting the laws of gamma processes. However, these studies give less attention to the statistical properties of the tilted processes. Our studies in fact focus on showing conjugate property of the random probability measure derived from tilted laws of the completely random measures and providing the posterior analysis of the class of random probability measures.

Applications of non-parametric models are becoming increasingly common in Bayesian statistics. However, implementing non-parametric models is rarely straightforward and often involves Markov chain Monte Carlo (MCMC) algorithms that might require evaluation of complicated functions. This article provides an augmented form of the sampling algorithm, namely the Blackwell–MacQueen Pólya urn sampling scheme, which avoids the necessity of evaluating those functions for some special cases and which we believe would be beneficial to the use of normalized random measures in statistical applications in future. We provide a simulation study concerned with estimating probabilities of partition sizes, a problem that arises in biological speciation (Lijoi *et al.* [33]). We show our algorithm yields similar results to three other sampling schemes based on the Blackwell–MacQueen Pólya urn distribution.

The article proceeds as follows. Section 2 describes the construction of a class of random probability measures obtained by tilting the laws of completely random measures. Materials presented here are in compact form; we refer the reader to Daley and Vere-Jones [8] for the complete treatment on these topics (see also Kingman [25–27]; Kallenberg [23,24]). Section 3 considers a class of random probability measures constructed through polynomially and exponentially tilting the laws of completely random measures. Section 3 also provides details on the prior and posterior distributions, proves the conjugacy property and describes the augmented Blackwell–MacQueen sampling scheme. Section 4 describes two specific examples of the tilting strategy, tilting the laws of the generalized gamma process and the generalized Dirichlet process is demonstrated. Section 5 presents the simulation study. Section 6 concludes the articles and provides a future research perspective. Proofs of theorems and propositions are included in the [Appendix](#).

In addition, as a referee pointed out that theorems in Section 3 are not entirely new and a version of them has appeared in an unpublished manuscript (James [18]), we would like to acknowledge that the class of random probability measures involving polynomially and exponentially tilting was first studied in the manuscript (James [18], Chapter 5), including a posterior analysis. We will further comments on these aspects and connections in Section 3.

2. Construction of the class of random probability measures

Let the triple $(\Omega, \mathfrak{F}, P)$ be the basic probability space. Assume \mathbb{X} is a Polish space endowed with a metric $d_{\mathbb{X}}$ generating a Borel σ -field $\mathcal{B}(\mathbb{X})$. Let $\mathcal{M}_{\mathbb{X}}$ denote the space of boundedly finite measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. A measure is said to be boundedly finite if it is finite on bounded sets. The space $\mathcal{M}_{\mathbb{X}}$ is a Polish space equipped with the metric of weak convergence. This induces the Borel σ -field $\mathcal{B}(\mathcal{M}_{\mathbb{X}})$. A *random measure*, μ say, taking values from $\mathcal{M}_{\mathbb{X}}$, is a measurable mapping $\mu : \Omega \times \mathbb{X} \rightarrow \mathbb{R}^+ \cup \{0\}$ where \mathbb{R}^+ denotes the positive real line. For each $\omega \in \Omega$, $\mu(\omega, \cdot)$ is a boundedly finite measure on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ and $\mu(\omega, A)$ is a random variable for all bounded sets $A \in \mathcal{B}(\mathbb{X})$. For convenience of notation, we write $\mu(A)$ instead of $\mu(\omega, A)$ from now on. For further details, see Daley and Vere-Jones [8], Chapter 9.

A random measure μ is a *completely random measure* (CRM) on the measure space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ if for all finite families of pairwise disjoint, bounded Borel sets $A_1, A_2, \dots, A_k \in \mathcal{B}(\mathbb{X})$, the random variables $\mu(A_1), \mu(A_2), \dots, \mu(A_k)$ are mutually independent. For any CRM, there is a representation theorem due to Kingman [25], Theorem 1 (see also Kingman [27] and Kingman [26], Chapter 8), and the theorem is nicely described in Daley and Vere-Jones [8], Theorem 10.1.III. The version in Daley and Vere-Jones [8], Theorem 10.1.III, says that a CRM μ can be represented as a sum of an atomic measure with countably many fixed atoms, a deterministic non-atomic measure and a measure derived from a Poisson process. The representation is given by

$$\mu(A) = \sum_{k=1}^{\infty} U_k \delta_{x_k}(A) + \lambda(A) + \int_{\mathbb{R}^+} y N(A, dy), \quad A \in \mathcal{B}(\mathbb{X}), \tag{2.1}$$

where the sequence $\{x_1, x_2, \dots\}$ is the countable set of fixed atoms of μ , $\{U_1, U_2, \dots\}$ is a sequence of mutually independent non-negative random variables, λ is a fixed non-atomic boundedly finite measure on \mathbb{X} , and N is a Poisson process on $\mathbb{X} \times \mathbb{R}^+$. This Poisson process N is independent of $\{U_1, U_2, \dots\}$ and has an intensity measure ν on $\mathbb{X} \times \mathbb{R}^+$. The intensity ν satisfies the following two conditions. For every bounded set $A \in \mathcal{B}(\mathbb{X})$,

$$\int_{\mathbb{R}^+} \min\{y, s\} \nu(A, dy) < \infty, \quad s \in \mathbb{R}^+, \tag{2.2}$$

and

$$\nu(\{x\}, \mathbb{R}^+) = 0, \quad x \in \mathbb{X}. \tag{2.3}$$

Notice that conditions (2.2) and (2.3) guarantee that the random measure $\int_{\mathbb{R}^+} y N(\cdot, dy)$ is boundedly finite on \mathbb{X} and has no fixed atoms respectively (see also Kallenberg [24], Chapter 12).

The Poisson process in (2.1) can be considered as a marked Poisson process on \mathbb{X} having mark space \mathbb{R}^+ . Again the product space $\mathbb{X} \times \mathbb{R}^+$ is a Polish space with a suitable metric $d_{\mathbb{X} \times \mathbb{R}^+}$ extended from $d_{\mathbb{X}}$. The Borel σ -field of the Polish space $\mathbb{X} \times \mathbb{R}^+$ is given by $\mathcal{B}(\mathbb{X} \times \mathbb{R}^+) = \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{R}^+)$ where $\mathcal{B}(\mathbb{R}^+)$ denotes the σ -field generated by the open subsets of \mathbb{R}^+ . Then the Poisson process N is a mapping $N : \Omega \times \mathbb{X} \times \mathbb{R}^+ \rightarrow \mathbb{Z}^+ \cup \{0\}$ where

\mathbb{Z}^+ denotes the positive integers. This Poisson process takes value from the space of boundedly finite measures $(\mathcal{M}_{\mathbb{X} \times \mathbb{R}^+}, \mathcal{B}(\mathcal{M}_{\mathbb{X} \times \mathbb{R}^+}))$ defined analogous to the space $(\mathcal{M}_{\mathbb{X}}, \mathcal{B}(\mathcal{M}_{\mathbb{X}}))$ discussed in the first paragraph. Here we write $N(A, B)$ to represent $N(\omega, A, B)$ for convenience of notation. The intensity measure ν of this Poisson process is a non-atomic σ -finite measure $\nu: \mathbb{X} \times \mathbb{R}^+ \rightarrow \mathbb{R}^+ \cup \{0\}$. The intensity measure is particularly important since it is the only parameter of the random measure and also the first moment of the measure. It determines the nature of the process and further it also determines the nature of the random measures derived from the Poisson process. Here (2.3) also implies that $N(\cdot, \mathbb{R}^+)$ is a simple point process on \mathbb{X} . This point process $N(\cdot, \mathbb{R}^+)$ is called ground process in Daley and Vere-Jones [8], Chapter 9.

Initially, we restrict our attention to the CRM on \mathbb{X} without the first two components, the atomic component and the drift, in (2.1), that is the CRM is in the form of

$$\tilde{\mu}(A) := \int_{\mathbb{R}^+} yN(A, dy), \quad A \in \mathcal{B}(\mathbb{X}), \tag{2.4}$$

with respect to the Poisson process N defined on $\mathbb{X} \times \mathbb{R}^+$ with intensity measure ν satisfying conditions (2.2) and (2.3). The law of the CRM $\tilde{\mu}$, denoted by $\mathcal{P}_{\tilde{\mu}}$, which is derived from the law of the Poisson process. For the sake of simplicity, we say $\tilde{\mu}$ has the parameter measure ν . The measure ν can be decomposed into two measures, ρ_x and η , written as $\nu(dx, ds) = \eta(dx)\rho_x(ds)$. Such a decomposition is guaranteed by Kallenberg [23], Appendix 15.3.3, in which the measure ρ_x is uniquely determined outside any set of ν measure zero. Here ρ_x is a mapping $\rho_x: \mathbb{R}^+ \rightarrow \mathbb{R}^+ \cup \{0\}$ for any $x \in \mathbb{X}$ such that $\rho_x(A)$ is \mathbb{X} measurable for every bounded set $A \in \mathcal{B}(\mathbb{R}^+)$ and ρ_x is a σ -finite measure. In particular, when ρ_x is dependent on $x \in \mathbb{X}$, the CRM $\tilde{\mu}$ is non-homogeneous. Otherwise, when ρ_x is not dependent on $x \in \mathbb{X}$, the CRM $\tilde{\mu}$ is homogeneous. The σ -finiteness of ρ_x ensures the CRM $\tilde{\mu}$ has countably infinite jumps on any bounded set in $\mathcal{B}(\mathbb{X})$. Here η is a finite non-atomic measure $\eta: \mathbb{X} \rightarrow \mathbb{R}^+ \cup \{0\}$. Without loss of generality, the measure η is restricted to be a proper probability measure on \mathbb{X} . This implies that the total mass of the measure $\tilde{\mu}$ is finite almost surely. Then a random probability measure could be defined according to the ratio of $\tilde{\mu}(A)$ and the total mass $\tilde{\mu}(\mathbb{X})$, that is $\tilde{G}(A) := \tilde{\mu}(A)/\tilde{\mu}(\mathbb{X})$ for $A \in \mathcal{B}(\mathbb{X})$.

Let h be a positive Borel measurable function $h: \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}^+$. Here $h(\tilde{\mu}(\mathbb{X}))$ is a tilting factor transforming the total mass to a positive scalar. The law of the *tilted completely random measure* (tilted CRM) is given by scaling the law of the CRM by the tilting factor. To ensure that the law of the tilted CRM is proper, the proportional constant of the law, $E[h(\tilde{\mu}(\mathbb{X}))]$, is required to be finite, that is,

$$E[h(\tilde{\mu}(\mathbb{X}))] < \infty. \tag{2.5}$$

Definition 2.1. Let $\tilde{\mu}$ be a CRM defined on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ in (2.4). The CRM $\tilde{\mu}$ has a probability measure $\mathcal{P}_{\tilde{\mu}}$ on $(\mathcal{M}_{\mathbb{X}}, \mathcal{B}(\mathcal{M}_{\mathbb{X}}))$ and with the parameter measure ν that satisfies conditions (2.2) and (2.3). Let h be a positive Borel measurable function on the non-negative real line, that satisfies condition (2.5). A tilted CRM, $\tilde{\mu}_t$, defined on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, has a probability measure $\mathcal{P}_{\tilde{\mu}_t}$ on $(\mathcal{M}_{\mathbb{X}}, \mathcal{B}(\mathcal{M}_{\mathbb{X}}))$ such that

$$\mathcal{P}_{\tilde{\mu}_t}(A) := \frac{1}{E[h(\tilde{\mu}(\mathbb{X}))]} \int_A h(\mu(\mathbb{X}))\mathcal{P}_{\tilde{\mu}}(d\mu), \quad A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}), \tag{2.6}$$

with the parameter measures ν and h .

Definition 2.2. Let $\tilde{\mu}_t$ be a tilted CRM with the parameter measures ν and h defined from Definition 2.1, a normalized tilted CRM \tilde{G}_t is given by

$$\tilde{G}_t(A) = \tilde{\mu}_t(A) / \tilde{\mu}_t(\mathbb{X}), \quad A \in \mathcal{B}(\mathbb{X}), \tag{2.7}$$

on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. This normalized tilted CRM \tilde{G}_t is with the parameter measures ν and h .

Notice that when the function is a finite constant ($h(x) = \text{constant} < \infty$), the normalized tilted CRM \tilde{G}_t is simply a normalized CRM, which has been extensively studied by James [18] and James *et al.* [22]. Some special cases with various choices of function h and CRMs have been considered. James [20] considers polynomially tilting the law of gamma process with $h(\mu(\mathbb{X})) = \mu(\mathbb{X})^{-q}$. Pitman [41,42] constructs the Poisson–Dirichlet process from polynomially tilting the law of positive stable process. These are all interesting special cases covered by the class of the normalized tilted CRMs which will be further discussed in Section 3.

The total mass is a key ingredient of both normalized CRMs and normalized tilted CRMs. We consider the connection between these two total masses and the general framework on a characterization of the masses through the Laplace transform. Let the total masses of CRM and tilted CRM be $\tilde{T} := \tilde{\mu}(\mathbb{X})$ and $\tilde{T}_t := \tilde{\mu}_t(\mathbb{X})$. Both $\tilde{\mu}$ and $\tilde{\mu}_t$ are the mappings to the positive line so that $\tilde{T} > 0$ and $\tilde{T}_t > 0$ and the laws of \tilde{T} and \tilde{T}_t are both absolutely continuous with respect to Lebesgue measure. Their densities, $f_{\tilde{T}}(y)$ and $f_{\tilde{T}_t}(y)$, are related through the equality

$$f_{\tilde{T}_t}(y) = \frac{h(y)f_{\tilde{T}}(y)}{\int_{\mathbb{R}^+} h(y)f_{\tilde{T}}(y) dy}. \tag{2.8}$$

The Laplace transform of the random variable \tilde{T} is given by

$$E[e^{-\lambda\tilde{T}}] = \int_{\mathbb{R}^+} e^{-\lambda y} f_{\tilde{T}}(y) dy = e^{-\psi_0(\lambda)},$$

where in general $\psi_a(b)$ is given by

$$\psi_a(b) = \int_{\mathbb{R}^+ \times \mathbb{X}} (1 - e^{-bs})e^{-as} \rho_x(ds)\eta(dx), \quad a \geq 0, b > 0. \tag{2.9}$$

In terms of the density of \tilde{T} , the Laplace transform of the random variable \tilde{T}_t could be seen as

$$E[e^{-\lambda\tilde{T}_t}] = \int_{\mathbb{R}^+} e^{-\lambda y} f_{\tilde{T}_t}(y) dy = \int_{\mathbb{R}^+} e^{-\lambda y} \frac{h(y)f_{\tilde{T}}(y)}{\int_{\mathbb{R}^+} h(s)f_{\tilde{T}}(s) ds} dy. \tag{2.10}$$

From (2.10), it requires to specify h to derive an explicit form. In fact, it is clear that the total masses are connected through the equality (2.8) and this could be utilized to derive the distributional results of the masses. A further extension on characterizing the random measures may concern the Laplace transform of their functionals, specifically linear functionals, that play an

important role in the studies of random measures. One could see it as a generalization from measuring a set (e.g., equation (2.4)) as an indicator function to measuring a class of functions. Let $BM^+(\mathbb{X})$ denote the class of positive Borel measurable functions mapping from \mathbb{X} to $\mathbb{R}^+ \cup \{0\}$ and all functions in this class vanish outside the bounded sets of \mathbb{X} . Let g be a function in $BM^+(\mathbb{X})$ and let the functional defined as $\tilde{\mu}(g) := \int_{\mathbb{X}} g(x)\tilde{\mu}(dx)$. This functional is also regarded as a Poisson functional since it has an expression with respect to the Poisson process on $\mathbb{X} \times \mathbb{R}^+$, that is $\tilde{\mu}(g) = \int_{\mathbb{R}^+ \times \mathbb{X}} g(x)yN(dx, dy)$. Then the properties of the CRM $\tilde{\mu}$ could be derived from the Poisson process N . For general discussion of functionals and Laplace functionals, see Daley and Vere-Jones [8], Chapters 9, 10. An early application to Bayesian non-parametric statistics would be found in [39] (see also Dykstra and Laud [11]).

3. The normalized random measure derived from the polynomially and exponentially tilted law

Here we aim at showing the conjugacy property of the normalized tilted CRM \tilde{G}_t with a specific choice of h (Definition 2.2). First, we take h as follows

$$h(x) = h'(x)x^{-q}, \quad q \geq 0, \tag{3.1}$$

where h' is a positive measurable function $h' : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}^+$ and satisfying condition (2.5), that is $E[h'(\tilde{\mu}(\mathbb{X}))\tilde{\mu}(\mathbb{X})^{-q}] < \infty$, but it is not depending on the scalar q . Take the normalized tilted CRM \tilde{G}_t be a prior random probability measure in the Bayesian non-parametric content, we can show that the posterior of \tilde{G}_t is belong to the same class of random probability measure (Definition 2.2). So the conjugacy property of \tilde{G}_t with the choice of h (3.1) is immediately revealed. We then consider the polynomially and exponentially tilted law of the CRM $\tilde{\mu}_t$, specifically, we take

$$h'(x) = e^{-\gamma x}, \quad \gamma \geq 0, \tag{3.2}$$

in (3.1). This choice covers a rich class of random probability measures and the interest in this choice is desirable. The posterior analysis of this class of tilted CRMs is given after the conjugacy property has been shown.

In general, the law of the tilted CRM $\tilde{\mu}_t$ is given by

$$\mathcal{P}_{\tilde{\mu}_t}(A) = \frac{1}{E[h'(\tilde{\mu}(\mathbb{X}))\tilde{\mu}(\mathbb{X})^{-q}]} \int_A h'(\mu(\mathbb{X}))\mu(\mathbb{X})^{-q} \mathcal{P}_{\tilde{\mu}}(d\mu), \quad A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}).$$

The parameters of both the tilted CRM $\tilde{\mu}_t$ and normalized tilted CRM \tilde{G}_t are now $q \geq 0$, h' and the intensity measure ν only. In particular, if h' is chosen to be $e^{-\gamma x}$ in (3.1), the parameters are then $q \geq 0$, $\gamma \geq 0$ and the intensity measure ν . The random probability measures in the class of normalized tilted CRMs are more general than those in the class of the normalized CRMs (Regazzini *et al.* [44]; James *et al.* [22]); one could easily realize that the tilted CRM $\tilde{\mu}_t$ (Definition 2.1) is not necessarily a CRM and this could be seen as a generalization of the class.

Remark 3.1. A referee pointed out that here the exponentially tilting, say $h(x) = e^{-\gamma x}$, is redundant as the exponentially tilting operation on a CRM leads to another CRM. So taking h' as in (3.2), such that $h(x) = e^{-\gamma x} x^{-q}$, operated with a CRM is equivalent to taking h' as a fixed finite constant, such that $h(x) = x^{-q}$, operated with another CRM.

Remark 3.2. Scaling operation on the law of Poisson random measures and CRMs has been studied in James [18]. James [18], equation 70, has the same construction as in this work by taking h' to be a fixed finite constant in (3.1), or to be (3.2) from Definition 2.1.

3.1. The posterior law and structural conjugacy

Consider a sequence of exchangeable random elements taking values in \mathbb{X} . These random variables are assumed to be conditionally independent and identically distributed given the normalized tilted CRM \tilde{G}_t (Definition 2.2) such that for every integer $n \geq 1$

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n | \tilde{G}_t) = \prod_{i=1}^n \tilde{G}_t(B_i) = \prod_{i=1}^n \frac{\tilde{\mu}_t(B_i)}{\tilde{\mu}_t(\mathbb{X})}, \quad B_1, \dots, B_n \in \mathcal{B}(\mathbb{X}), \quad (3.3)$$

where \tilde{G}_t , or equivalently $\tilde{\mu}_t$, is regarded as a parameter. Then, (3.3) is the ‘‘likelihood’’ with the parameter $\tilde{\mu}_t$. Let $\mathcal{P}_{\tilde{\mu}_t}^{(\mathbf{X}_n)}$ indicate the posterior distribution of $\tilde{\mu}_t$, namely the distribution of $\tilde{\mu}_t$ conditional on \mathbf{X}_n , that is

$$\mathcal{P}_{\tilde{\mu}_t}^{(\mathbf{X}_n)}(A) := \frac{\mathcal{P}_{\tilde{\mu}_t, \mathbf{X}_n}(A, B_1, \dots, B_n)}{\mathcal{P}_{\tilde{\mu}_t, \mathbf{X}_n}(\mathcal{M}_{\mathbb{X}}, B_1, \dots, B_n)}, \quad A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}), B_1, \dots, B_n \in \mathcal{B}(\mathbb{X}),$$

where $\mathcal{P}_{\tilde{\mu}_t, \mathbf{X}_n}$ represents the joint distribution of $(\tilde{\mu}_t, \mathbf{X}_n)$ such that

$$\mathcal{P}_{\tilde{\mu}_t, \mathbf{X}_n}(A, B_1, \dots, B_n) = \int_A \prod_{i=1}^n \frac{\mu(B_i)}{\mu(\mathbb{X})} \mathcal{P}_{\tilde{\mu}_t}(d\mu), \quad A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}), B_1, \dots, B_n \in \mathcal{B}(\mathbb{X}). \quad (3.4)$$

Taking into consideration of the above assumption of conditional independence (3.3), the joint distribution $\mathcal{P}_{\tilde{\mu}_t, \mathbf{X}_n}$ of $(\tilde{\mu}_t, \mathbf{X}_n)$ is defined on the space $(\mathcal{M}_{\mathbb{X}} \times \mathbb{X}^n, \mathcal{B}(\mathcal{M}_{\mathbb{X}}) \otimes \mathcal{B}(\mathbb{X})^n)$ for $n = 1, 2, \dots$. Now, considering Definition 2.1 of $\mathcal{P}_{\tilde{\mu}_t}$, one obtains

$$\int_A \prod_{i=1}^n \frac{\mu(B_i)}{\mu(\mathbb{X})} \mathcal{P}_{\tilde{\mu}_t}(d\mu) = \frac{1}{E[h'(\tilde{\mu}(\mathbb{X}))\tilde{\mu}(\mathbb{X})^{-q}]} \int_A \frac{h'(\mu(\mathbb{X}))}{\mu(\mathbb{X})^{n+q}} \left(\prod_{i=1}^n \mu(B_i) \right) \mathcal{P}_{\tilde{\mu}}(d\mu), \quad (3.5)$$

$$A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}), B_1, \dots, B_n \in \mathcal{B}(\mathbb{X}).$$

This shows the key element needed for deriving the posterior law of $\tilde{\mu}_t$ is the law of the CRM $\mathcal{P}_{\tilde{\mu}}$ as seen in the right-hand side of (3.5). The usual technique to derive the posterior of $\tilde{\mu}_t$ requires application of change of measure or disintegration. So, the major task is to apply change of measure updating the law $\mathcal{P}_{\tilde{\mu}}$ with the information, namely $(\prod_{i=1}^n \mu(B_i))$, $h'(\mu(\mathbb{X}))$ and

$\mu(\mathbb{X})^{-(n+q)}$, to a posterior law, $\mathcal{P}_{\tilde{\mu}}^{\mathbb{X}^n}$. Dealing with the term $(\prod_{i=1}^n \mu(B_i))$ could be simply adopted by the standard arguments (James [18–20]). The next term $h'(\mu(\mathbb{X}))$ should be chosen explicitly to proceed. When $h'(\mu(\mathbb{X})) = e^{-\gamma\mu(\mathbb{X})}$, change of measure only involves the Laplace transform and doesn't cost much effort. Eventually, dealing with the term $\mu(\mathbb{X})^{-(n+q)}$ could be somehow challenging. James [18] (see also James *et al.* [22]) introduces an augmentation approach that allows us to proceed further in particular for this polynomial term $\mu(\mathbb{X})^{-(n+q)}$. This leads to the analysis on the posteriors of the normalized tilted CRMs.

James [18]'s approach makes use of the gamma identity and introduces an augmented variable. We now address the role of the augmented variable. Here the well-known gamma identity is given by

$$\frac{1}{b^a} = \frac{1}{\Gamma(a)} \int_{\mathbb{R}^+} e^{-bu} u^{a-1} du, \quad a, b \in \mathbb{R}^+. \tag{3.6}$$

Then take the term $\mu(\mathbb{X})^{-(n+q)}$ in (3.5) as b^{-a} in (3.6), term involving μ becomes tractable, positioned as an exponent. The integral (numerator) appearing in the right-hand side (3.5) can be rewritten as

$$\int_{A \times \mathbb{R}^+} h'(y) \left[\int_{\mathbb{R}^+} \frac{u^{n+q-1} e^{-uy}}{\Gamma(n+q)} du \right] \left(\prod_{i=1}^n \mu(B_i) \right) \mathbb{P}(\tilde{\mu} \in d\mu | \tilde{T} = y) f_{\tilde{T}}(y) dy, \tag{3.7}$$

$A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}), B_1, \dots, B_n \in \mathcal{B}(\mathbb{X}).$

where $\mathbb{P}(\tilde{\mu} \in d\mu | \tilde{T} = y)$ denotes the conditional distribution of $\tilde{\mu}$ given its' total mass $\tilde{T} = \tilde{\mu}(\mathbb{X})$. Rewriting the last expression by replacing $f_{\tilde{T}}$ with its expression in term $f_{\tilde{T}_i}$ (2.8), from (3.5) and (3.7) one finally gets

$$\int_A \prod_{i=1}^n \frac{\mu(B_i)}{\mu(\mathbb{X})} \mathcal{P}_{\tilde{\mu}_i}(d\mu) = \int_{A \times \mathbb{R}^+ \times \mathbb{R}^+} \left(\prod_{i=1}^n \frac{\mu(B_i)}{y} \right) \mathbb{P}(\tilde{\mu} \in d\mu | \tilde{T} = y) \ell_n(y, u) dy du, \tag{3.8}$$

$A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}), B_1, \dots, B_n \in \mathcal{B}(\mathbb{X}).$

where

$$\ell_n(y, u) = \frac{u^{n+q-1} e^{-uy} y^{n+q}}{\Gamma(n+q)} f_{\tilde{T}_i}(y), \quad u > 0, y > 0. \tag{3.9}$$

Here ℓ_n is a joint probability density on $\mathbb{R}^+ \times \mathbb{R}^+$. Without loss of generality, we assume that $(\Omega, \mathfrak{F}, \mathbb{P})$ is large enough to support a sequence U_n of random numbers such that the distribution of (U_n, \tilde{T}_i) admits ℓ_n as density function. Then, it follows that

$$f_{U_n}(u) = \int_{\mathbb{R}^+} \ell_n(y, u) dy = \frac{u^{n+q-1}}{\Gamma(n+q)} \int_{\mathbb{R}^+} y^{n+q} e^{-uy} f_{\tilde{T}_i}(y) dy, \quad u > 0. \tag{3.10}$$

Here f_{U_n} is a probability density for U_n . In view of these elementary developments, one can disintegrate the law of $(\tilde{\mu}_t, \mathbf{X}_n)$ as follows

$$\int_A \prod_{i=1}^n \frac{\mu(B_i)}{\mu(\mathbb{X})} \mathcal{P}_{\tilde{\mu}_t}(\mathrm{d}\mu) = \int_{A \times \mathbb{R}^+ \times \mathbb{R}^+} \left(\prod_{i=1}^n \frac{\mu(B_i)}{y} \right) \times \mathbb{P}(\tilde{\mu} \in \mathrm{d}\mu | \tilde{T} = y) f_{\tilde{T}|U_n}(y|u) f_{U_n}(u) \mathrm{d}u \mathrm{d}y, \quad (3.11)$$

$$A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}}), B_1, \dots, B_n \in \mathcal{B}(\mathbb{X}),$$

where

$$f_{\tilde{T}|U_n}(y|u) = \frac{\ell_n(u, y)}{\int_{\mathbb{R}^+} \ell_n(u, y) \mathrm{d}y} = \frac{y^{n+q} e^{-uy}}{E[\tilde{T}_t^{n+q} e^{-u\tilde{T}_t}]} f_{\tilde{T}_t}(y), \quad y > 0.$$

This disintegration (3.11) shows the role of U_n as an augmented variable.

Remark 3.3. The representation of the density (3.10) suggests that U_n has the gamma distribution with a random scale \tilde{T}_t which has the density $f_{\tilde{T}_t}$. Given \tilde{T}_t , U_n has the gamma distribution with parameters $(n + q, \tilde{T}_t)$. The product of the random variables, $U_n \tilde{T}_t$, has the gamma distribution with parameters $(n + q, 1)$ and independent of \tilde{T}_t .

Before proceeding to the posterior distribution, additional notations are introduced. The normalized tilted CRM \tilde{G}_t is almost surely discrete. A random sample \mathbf{X}_n of \tilde{G}_t usually contains ties. We can always express \mathbf{X}_n by two elements, namely a partition and distinct values. Here \mathbf{p}_n is a partition of the integers $\{1, \dots, n\}$ that are the indices of \mathbf{X}_n and $\mathbf{Y}_{\mathfrak{N}(\mathbf{p}_n)} = \{Y_1, \dots, Y_{\mathfrak{N}(\mathbf{p}_n)}\}$ represents the distinct values of \mathbf{X}_n . The partition \mathbf{p}_n locates the distinct values from \mathbf{X}_n to $\mathbf{Y}_{\mathfrak{N}(\mathbf{p}_n)}$ or vice versa. As a result, we have the following equivalent representations

$$\mathbf{X}_n = \{X_1, \dots, X_n\} = \{Y_1, \dots, Y_{\mathfrak{N}(\mathbf{p}_n)}, \mathbf{p}_n\} = \{\mathbf{Y}_{\mathfrak{N}(\mathbf{p}_n)}, \mathbf{p}_n\}. \quad (3.12)$$

A partition \mathbf{p}_n contains $\mathfrak{N}(\mathbf{p}_n)$ cells (known as clusters), that is $\mathbf{p}_n = \{C_1, \dots, C_{\mathfrak{N}(\mathbf{p}_n)}\}$. Each cell C_k contains the indices of a subset of \mathbf{X}_n , namely the unique values Y_k such that $C_k = \{i : X_i = Y_k, i = 1, \dots, n\}$ for $k = 1, \dots, \mathfrak{N}(\mathbf{p}_n)$. The number of elements in the cell k , C_k , of the partition is indicated by n_k , for $k = 1, \dots, \mathfrak{N}(\mathbf{p}_n)$, so that $\sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} n_k = n$. Therefore, the union of all cells is the set of all n integers, $\bigcup_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} C_k = \{1, \dots, n\}$ and all cells are pairwise mutually exclusive, that is $C_k \cap C_{k'} = \emptyset$ where $k \neq k'$ for $k, k' = 1, \dots, \mathfrak{N}(\mathbf{p}_n)$. This partition representation is commonly used in Bayesian non-parametric literature (see Lo [36]; Lo and Weng [39]; James [18–20]) since it well describes the variates generated from those random probability measures and is also useful in expressing the marginal distribution of \mathbf{X}_n .

Lijoi and Prünster [35] describe the concept of structural conjugacy. A random probability measure, say G , is a structurally conjugate random probability measure if the resulting posterior law of G given $\mathbf{X}_n = \{X_1, \dots, X_n\}$, has the same structure. *Neutral to the right* process (Doksum [10]) is one of the classes that has this property. In the present work, we show that the normalized tilted CRMs in Definition 2.1 are also structurally conjugate. Here we follow James *et al.* [22],

write $\tilde{G}_t^{(\mathbf{X}_n)}$ as the posterior normalized tilted CRM. Theorem 3.1 shows that the normalized tilted CRMs in Definition 2.2 have the conjugate property, that is, both \tilde{G}_t and $\tilde{G}_t^{(\mathbf{X}_n)}$ are in the same class.

Theorem 3.1. *Let $\tilde{G}_t = \tilde{\mu}_t / \tilde{\mu}_t(\mathbb{X})$ be a normalized tilted CRM (Definition 2.2) defined on \mathbb{X} with h as in (3.1), that is $h(x) = h'(x)x^{-q}$. The parameters of this normalized tilted CRM are given by the Borel measurable function h' , the measure ν , and the scalar $q \geq 0$. With the prior measure \tilde{G}_t and the likelihood (3.3), the posterior measure, namely $\tilde{G}_t^{(\mathbf{X}_n)}$, has the same distribution of a normalized tilted CRM such that*

$$\tilde{G}_t^{(\mathbf{X}_n)} \stackrel{d}{=} \tilde{\mu}_t^{(\mathbf{X}_n)} / \tilde{\mu}_t^{(\mathbf{X}_n)}(\mathbb{X}), \tag{3.13}$$

where

1. $\tilde{\mu}_t^{(\mathbf{X}_n)}$ has the law

$$\mathcal{P}_{\tilde{\mu}_t^{(\mathbf{X}_n)}}(A) = \frac{1}{E[h'(\tilde{\mu}^{(\mathbf{X}_n)}(\mathbb{X}))\tilde{\mu}^{(\mathbf{X}_n)}(\mathbb{X})^{-(n+q)}]} \int_A h'(\mu(\mathbb{X}))\mu(\mathbb{X})^{-(n+q)}\mathcal{P}_{\tilde{\mu}^{(\mathbf{X}_n)}}(d\mu)$$

for $A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}})$.

2. $\tilde{\mu}^{(\mathbf{X}_n)}$ is a normalized tilted CRM whose law $\mathcal{P}_{\tilde{\mu}^{(\mathbf{X}_n)}}$, has the same law as

$$\tilde{\mu} + \sum_{k=1}^{\aleph(\mathbf{p}_n)} J_k^{(\mathbf{X}_n)} \delta_{Y_k}, \tag{3.14}$$

where δ_a is the Dirac delta function evaluated at a , $\tilde{\mu}$ is a CRM with law $\mathcal{P}_{\tilde{\mu}}$ and intensity measure ν , $\{Y_1, \dots, Y_{\aleph(\mathbf{p}_n)}\}$ is a sequence of fixed points of discontinuity, and $\{J_1^{(\mathbf{X}_n)}, \dots, J_{\aleph(\mathbf{p}_n)}^{(\mathbf{X}_n)}\}$ are the corresponding jumps.

3. Conditional on \mathbf{X}_n , each jump $J_k^{(\mathbf{X}_n)}$ has the conditional distribution

$$\mathbb{P}(J_k^{(\mathbf{X}_n)} \in ds) = \frac{s^{n_k} \rho_{Y_k}(ds)}{\int_{\mathbb{R}^+} s^{n_k} \rho_{Y_k}(ds)}, \quad k = 1, \dots, \aleph(\mathbf{p}_n).$$

4. Conditional on \mathbf{X}_n , $\tilde{\mu}$ and $\{J_1^{(\mathbf{X}_n)}, \dots, J_{\aleph(\mathbf{p}_n)}^{(\mathbf{X}_n)}\}$ are independent.

Proof. See Appendix A.1. □

In Theorem 3.1, the posterior law of $\tilde{\mu}_t$, namely $\tilde{\mu}_t^{(\mathbf{X}_n)}$, is shown to have the tilted law of the CRM $\tilde{\mu}^{(\mathbf{X}_n)}$, which has the law $\mathcal{P}_{\tilde{\mu}^{(\mathbf{X}_n)}}$, where the tilting factor is updated to $h'(\tilde{\mu}^{(\mathbf{X}_n)}(\mathbb{X}))\tilde{\mu}^{(\mathbf{X}_n)}(\mathbb{X})^{-(n+q)}$. Under the normalization of the process, the posterior of the random probability measure becomes $\tilde{G}_t^{(\mathbf{X}_n)} \stackrel{d}{=} \tilde{\mu}_t^{(\mathbf{X}_n)} / \tilde{\mu}_t^{(\mathbf{X}_n)}(\mathbb{X})$. This confirms the conjugate property of the normalized tilted CRM.

Remark 3.4. Following Remark 3.2, the posterior distribution discussed in James [18], Chapter 5, over the scaling operation on CRMs has been established in James [18], Corollary 5.1. This result is a version of Theorem 3.1. Furthermore, James [18], Theorem 5.1, also supplies the posterior law of the corresponding Poisson random measure. In particular, putting statements i and ii of James [18], Corollary 5.1, and statements i and ii of James [18], Proposition 5.2, together is equivalent to Theorem 3.1.

A special case of the tilted CRM (Definition 2.1) that takes $h(x) = e^{-\gamma x} x^{-q}$ is of interest. The following theorem, Theorem 3.2, describes an augmented posterior law of \tilde{G}_t and $\tilde{\mu}_t$, denoted by $\tilde{G}_t^{(U_n, \mathbf{X}_n)}$ and $\tilde{\mu}_t^{(U_n, \mathbf{X}_n)}$, that are the conditional laws of \tilde{G}_t and $\tilde{\mu}_t$, respectively, given both U_n and \mathbf{X}_n .

Theorem 3.2. Let $\tilde{G}_t = \tilde{\mu}_t / \tilde{\mu}_t(\mathbb{X})$ be a normalized tilted CRM (Definition 2.2) on \mathbb{X} with h as in (3.1) and h' as in (3.2), that is $h(x) = e^{-\gamma x} x^{-q}$. The parameters of this normalized tilted CRM are given by the measure ν , and the scalars $q \geq 0$ and $\gamma \geq 0$. With the prior measure \tilde{G}_t and the likelihood (3.3), the conditional posterior measure, namely $\tilde{G}_t^{(U_n, \mathbf{X}_n)}$, given U_n and $\mathbf{X}_n = \{X_1, \dots, X_n\}$ has the same distribution of a normalized CRM such that

$$\tilde{G}_t^{(U_n, \mathbf{X}_n)} \stackrel{d}{=} \tilde{\mu}^{(U_n, \mathbf{X}_n)} / \tilde{\mu}^{(U_n, \mathbf{X}_n)}(\mathbb{X}),$$

where

1. $\tilde{\mu}^{(U_n, \mathbf{X}_n)}$ is a CRM with the same law as

$$\tilde{\mu}^{(U_n)} + \sum_{k=1}^{\aleph(\mathbf{p}_n)} J_k^{(U_n, \mathbf{X}_n)} \delta_{Y_k}, \tag{3.15}$$

where $\tilde{\mu}^{(U_n)}$ is a CRM with intensity $\nu(dx, ds) = \eta(dx) \times e^{-s(\gamma+U_n)} \rho_x(ds)$, $\{Y_1, \dots, Y_{\aleph(\mathbf{p}_n)}\}$ is a sequence of fixed points of discontinuity, and $\{J_1^{(U_n, \mathbf{X}_n)}, \dots, J_{\aleph(\mathbf{p}_n)}^{(U_n, \mathbf{X}_n)}\}$ are the corresponding jumps.

2. Conditional on U_n and \mathbf{X}_n , each jump $J_k^{(U_n, \mathbf{X}_n)}$ has the conditional density

$$\mathbb{P}(J_k^{(U_n, \mathbf{X}_n)} \in ds) = \frac{s^{n_k} e^{-s(\gamma+U_n)} \rho_{Y_k}(ds)}{\int_{\mathbb{R}^+} s^{n_k} e^{-s(\gamma+U_n)} \rho_{Y_k}(ds)}, \quad k = 1, \dots, \aleph(\mathbf{p}_n).$$

3. Conditional on U_n and \mathbf{X}_n , $\mu^{(U_n)}$ and $\{J_1^{(U_n, \mathbf{X}_n)}, \dots, J_{\aleph(\mathbf{p}_n)}^{(U_n, \mathbf{X}_n)}\}$ are independent.
4. Conditional on \mathbf{X}_n , U_n has the density

$$f_{U_n|\mathbf{X}_n}(u) = \frac{e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \tau_{n_k, Y_k}(\gamma+u) u^{n+q-1}}{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k'=1}^{\aleph(\mathbf{p}_n)} \tau_{n_{k'}, Y_{k'}}(\gamma+u) u^{n+q-1} du}, \tag{3.16}$$

where

$$\tau_{m,z}(a) = \int_{\mathbb{R}^+} s^m e^{-sa} \rho_z(ds), \quad m > 0, z \in \mathbb{X}, a \geq 0, \tag{3.17}$$

and $\psi_a(b)$ for $a \geq 0$ and $b \geq 0$ is defined in (2.9).

Proof. See Appendix A.2. □

Theorem 3.2 shows that the augmented posterior random probability measure $\tilde{G}_t^{(U_n, \mathbf{X}_n)}$ has the same distribution of a normalized CRM and further also provides the posterior distribution of the augmented variables $U_n, f_{U_n|\mathbf{X}_n}$. Combining these two yields the posterior random probability measure $\tilde{G}_t^{(\mathbf{X}_n)}$. This is certainly useful in applications of Bayesian non-parametric. For example, this theorem could be useful in simulating the posterior normalized tilted CRM $\tilde{G}_t^{(\mathbf{X}_n)}$ that is desirable in some applications.

Remark 3.5. The posterior distribution of the tilted CRM, namely $\tilde{\mu}_t^{(\mathbf{X}_n)}$, can be achieved by mixing the law in (3.15) (Theorem 3.2, statement 1) over the distribution of U_n given \mathbf{X}_n in (3.16) (Theorem 3.2, statement 4). This is equivalent to the law of the tilted CRM in statement i of James [18], Corollary 5.1.

3.2. Generalized Blackwell and MacQueen Pólya urn sampling scheme and marginal distribution of partitions

Blackwell and MacQueen [3] first introduce the Pólya urn sampling scheme for the Dirichlet process and this scheme can be utilized to generate random sequences from the Dirichlet process. This employs so called the *Blackwell–MacQueen Pólya* (BMP) urn formula, that is the predictive distribution of the Dirichlet process random sequences. Lo [37] also shows that this BMP urn formula can be used to characterize the Dirichlet process. James *et al.* [22] generalize the Blackwell–MacQueen Pólya (BMP) urn formula for the normalized CRMs, namely *Generalized Blackwell–MacQueen Pólya* (GBMP) urn formula. A further generalization will be considered for the normalized tilted CRM in this section.

We consider a normalized tilted CRM \tilde{G}_t with $h(x) = e^{-\gamma x} x^{-q}$ in Definition 2.2. Here the GBMP urn formula for this normalized tilted CRM will be presented under two formulations, namely the unconditional and the conditional GBMP urn formulas for the normalized tilted CRM. The unconditional GBMP urn formula is simply the predictive distribution, $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\}$ where $\{X_{n+1}, \mathbf{X}_n\}$ is the random sequence drawn from the normalized tilted CRM \tilde{G}_t . The conditional GBMP urn formula is the augmented version of GBMP urn formula derived from Theorem 3.2. An impression directly comes to the mind is that the conditional urn formula, namely $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, U_n\}$ could be derived according to the predictive distribution $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\} = E[\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, U_n\} | \mathbf{X}_n]$ with respect to the distribution U_n given $\mathbf{X}_n, f_{U_n|\mathbf{X}_n}$ (3.16). However, the term $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, U_n\}$ is not necessarily a proper distribution. A rescaling over both $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, U_n\}$ and $f_{U_n|\mathbf{X}_n}$ with a factor involving terms like U_n seems to be needed. This leads to a new variable \tilde{U}_n introduced and this new variable has a tilted density of U_n . So, we obtain a proper distribution $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\}$ and also the conditional GBMP urn formula is established by mixing over \tilde{U}_n .

Proposition 3.1. Let $\tilde{G}_t = \tilde{\mu}_t / \tilde{\mu}_t(\mathbb{X})$ be a normalized tilted CRM (Definition 2.2) defined on \mathbb{X} with h as in (3.1) and h' as in (3.2), that is $h(x) = e^{-\gamma x} x^{-q}$. The parameters of this normalized tilted CRM are given by the measure ν , and the scalars $q \geq 0$ and $\gamma \geq 0$. Then,

1. The predictive distribution for X_{n+1} given $\mathbf{X}_n = \{X_1, \dots, X_n\}$ is given by

$$\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\} = \frac{\omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(x)}{\phi(\mathbf{X}_n)} \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{\omega_{n,k}(Y_k)}{\phi(\mathbf{X}_n)} \delta_{Y_k}(dx), \tag{3.18}$$

where

$$\begin{aligned} \phi(\mathbf{X}_n) &= \int_{\mathbb{X}} \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(x) \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \omega_{n,k}(Y_k), \\ \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(x) &= \int_{\mathbb{R}^+} u \tau_{1,x}(\gamma + u) f_{U_n | \mathbf{X}_n}(u) du \quad \text{and} \\ \omega_{n,k}(Y_k) &= \int_{\mathbb{R}^+} u \frac{\tau_{n_k+1, Y_k}(\gamma + u)}{\tau_{n_k, Y_k}(\gamma + u)} f_{U_n | \mathbf{X}_n}(u) du, \quad k = 1, \dots, \mathfrak{N}(\mathbf{p}_n). \end{aligned}$$

2. Conditional on \tilde{U}_n , the predictive distribution for X_{n+1} given $\mathbf{X}_n = \{X_1, \dots, X_n\}$ is given by

$$\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} = \frac{\omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(\tilde{U}_n, x)}{\phi(\tilde{U}_n, \mathbf{X}_n)} \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{\omega_{n,k}(\tilde{U}_n, Y_k)}{\phi(\tilde{U}_n, \mathbf{X}_n)} \delta_{Y_k}(dx), \tag{3.19}$$

where

$$\begin{aligned} \phi(u, \mathbf{X}_n) &= \int_{\mathbb{X}} \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(u, x) \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \omega_{n,k}(u, Y_k), \\ \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(u, x) &= u \tau_{1,x}(\gamma + u), \\ \omega_{n,k}(u, Y_k) &= u \frac{\tau_{n_k+1, Y_k}(\gamma + u)}{\tau_{n_k, Y_k}(\gamma + u)}, \quad k = 1, \dots, \mathfrak{N}(\mathbf{p}_n) \quad \text{and} \\ f_{\tilde{U}_n | \mathbf{X}_n}(u) &= \frac{\phi(u, \mathbf{X}_n)}{E[\phi(U_n, \mathbf{X}_n) | \mathbf{X}_n]} f_{U_n | \mathbf{X}_n}(u), \quad u > 0. \end{aligned}$$

3. In addition,

$$\begin{aligned} \phi(\mathbf{X}_n) &= E[\phi(U_n, \mathbf{X}_n) | \mathbf{X}_n] = n + q, \\ \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(x) &= E[\omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(U_n, x) | \mathbf{X}_n] \quad \text{and} \\ \omega_{n,k}(Y_k) &= E[\omega_{n,k}(U_n, Y_k) | \mathbf{X}_n], \quad k = 1, \dots, \mathfrak{N}(\mathbf{p}_n), \end{aligned}$$

where $\psi_a(b)$ for $a \geq 0$ and $b \geq 0$ defined in (2.9) and $\tau_{m,z}(a)$ for $m > 0, z \in \mathbb{X}$ and $a \geq 0$ defined in (3.17).

Proof. See Appendix A.3. □

Proposition 3.1 gives the predictive distribution and the augmented predictive distribution for the sequence $\mathbf{X}_n = \{X_1, \dots, X_n\}$ as in Blackwell and MacQueen [3] and James *et al.* [22]. Statement 1 of Proposition 3.1 could be viewed as a direct sampling scheme and statement 2 as a conditional sampling scheme that involves iterative sampling with \tilde{U}_n given \mathbf{X}_n . From statement 1, conditional on $\mathbf{X}_n = \{\mathbf{Y}_{\aleph(\mathbf{p}_n)}, \mathbf{p}_n\}$, X_{n+1} is sampled from $[\int_{\mathbb{X}} \omega_{n,\aleph(\mathbf{p}_n)+1}(x)\eta(dx)]^{-1} \omega_{n,\aleph(\mathbf{p}_n)+1}(x)\eta(dx)$ with probability $[\phi(\mathbf{X}_n)]^{-1} \int_{\mathbb{X}} \omega_{n,\aleph(\mathbf{p}_n)+1}(x) \times \eta(dx)$ and the new sample is allocated a new index $\aleph(\mathbf{p}_n) + 1$, becoming $Y_{\aleph(\mathbf{p}_n)+1}$. Otherwise X_{n+1} has probability $1 - [\phi(\mathbf{X}_n)]^{-1} \int_{\mathbb{X}} \omega_{n,\aleph(\mathbf{p}_n)+1}(x)\eta(dx)$ to be one of the existing sample $\mathbf{Y}_{\aleph(\mathbf{p}_n)}$. This sequential scheme directly collects sample of size n cumulatively. Statement 2 of Proposition 3.1 suggests an alternative sampling scheme that draws from $\{X_i, \tilde{U}_i\}$ sequentially for $i = 1, \dots, n$. To initialize the scheme, a sample of \tilde{U}_0 from the marginal distribution $f_{\tilde{U}_0} = f_{U_0}$ and a sample of X_1 from $[\int_{\mathbb{X}} \tau_{1,x}(\gamma + \tilde{U}_0)\eta(dx)]^{-1} \tau_{1,x}(\gamma + \tilde{U}_0)\eta(dx)$ are required, then a sample of $\{X_2, \dots, X_n\}$ can be achieved through iterating the following steps

- Step 1: conditional on \mathbf{X}_i , sample \tilde{U}_i from $f_{\tilde{U}_i|\mathbf{X}_i}$,
- Step 2: conditional on \mathbf{X}_i and \tilde{U}_i , sample X_{i+1} from $\mathbb{P}\{X_{i+1} \in dx|\mathbf{X}_i, \tilde{U}_i\}$,

for $i = 1, \dots, n - 1$, where $\mathbf{X}_i = \{\mathbf{Y}_{\aleph(\mathbf{p}_i)}, \mathbf{p}_i\}$. The scheme described here is more general than those in existing articles and provides alternatives to sample from some common processes, such as the normalized generalized gamma process and the generalized Dirichlet process (see James [18]; Lijoi *et al.* [31,32] for a direct sampling scheme). These two cases will be discussed in Section 4.

Remark 3.6. The joint distribution of \mathbf{X}_n can be recovered according to (3.18) that can be also found in James [18], Theorem 5.2.

Proposition 3.2. Let $\tilde{G}_t = \tilde{\mu}_t / \tilde{\mu}_t(\mathbb{X})$ be a normalized tilted CRM (Definition 2.2) defined on \mathbb{X} with h as in (3.1) and h' as in (3.2), that is $h(x) = e^{-\gamma x} x^{-q}$. The parameters of this normalized tilted CRM are given by the measure ν , and the scalars $q \geq 0$ and $\gamma \geq 0$. Then,

1. The marginal distribution for $\mathbf{X}_n = \{\mathbf{Y}_{\aleph(\mathbf{p}_n)}, \mathbf{p}_n\}$ is given by

$$\frac{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \tau_{n_k, Y_k}(\gamma + u) u^{n+q-1} du}{\sum_{\mathbf{p}_n} \int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma + u) u^{n+q-1} du} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k),$$

where

$$\kappa_m(a) = \int_{\mathbb{X}} \tau_{m,x}(a)\eta(dx) = \int_{\mathbb{X}} \left[\int_{\mathbb{R}^+} s^m e^{-sa} \rho_x(ds) \right] \eta(dx), \quad m > 0, a \geq 0.$$

2. The marginal distribution for \mathbf{p}_n is given by

$$\frac{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma+u) u^{n+q-1} du}{\sum_{\mathbf{p}_n} \int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma+u) u^{n+q-1} du}.$$

3. Conditional on U_n , the distribution for $\mathbf{X}_n = \{\mathbf{Y}_{\aleph(\mathbf{p}_n)}, \mathbf{p}_n\}$ is given by

$$\frac{\prod_{k=1}^{\aleph(\mathbf{p}_n)} \tau_{n_k, Y_k}(\gamma + U_n)}{\sum_{\mathbf{p}_n} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma + U_n)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k).$$

4. Conditional on U_n , the distribution for \mathbf{p}_n is given by

$$\frac{\prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma + U_n)}{\sum_{\mathbf{p}_n} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma + U_n)}.$$

5. Conditional on \mathbf{p}_n , the distribution for U_n is given by

$$\frac{e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma+u) u^{n+q-1}}{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \kappa_{n_k}(\gamma+u) u^{n+q-1} du},$$

where $\psi_a(b)$ for $a \geq 0$ and $b \geq 0$ defined in (2.9) and $\tau_{m,z}(a)$ for $m > 0, z \in \mathbb{X}$ and $a \geq 0$ defined in (3.17).

Proof. See Appendix A.4. □

Proposition 3.2 gives the marginal distributions of \mathbf{X}_n , both conditional and unconditional on U_n , in statements 1 and 3, respectively. In statements 2 and 4, the proposition gives the distributions of \mathbf{p}_n of first n integers $\{1, \dots, n\}$ both conditional and unconditional on U_n . The distributions of the partitions are the *exchangeable partition probability function* (EPPF) as they all are symmetric functions of $\{n_1, \dots, n_{\aleph(\mathbf{p}_n)}\}$. A special structure of the EPPF, called the Gibbs form, is also available in the conditional case. Specifically, the distribution of the partition, \mathbf{p}_n , conditional on U_n (Statement 4 of Proposition 3.2) has the Gibbs form (Pitman [42], Theorem 4.6, page 86), that is, the EPPF is of the form $V_{n, \aleph(\mathbf{p}_n)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} W_{n_k}$ in which $V_{n, \aleph(\mathbf{p}_n)}$ is a function of n and $\aleph(\mathbf{p}_n)$ and W_{n_k} is a function of n_k . Such a partition is called the Gibbs partition and therefore \mathbf{p}_n is a Gibbs partition conditional on U_n (see Pitman [40] and Pitman [42] for the details of the EPPF and the Gibbs form).

Remark 3.7. The results appear in Proposition 3.2 can be derived from James [18], Theorem 5.2. Some special cases could be found in James [18], Sections 5.3 and 5.4.

4. The tilted version of generalized gamma process and generalized Dirichlet process

We consider the tilted version of two interesting and important classes of measures, namely the generalized gamma process and the generalized Dirichlet process. Tilting these two classes of measures yields the normalized beta-gamma process, the Poisson–Dirichlet process, the normalized generalized gamma process and the Dirichlet process. They all have been extensively studied and have wide applications in both statistics and probability. We now discuss the tilted version of these processes. Let \tilde{G}_t be a normalized tilted CRM with $h(x) = e^{-\gamma x} x^{-q}$ and with ν chosen in the following subsections.

4.1. Generalized gamma process

The generalized gamma process is considered a building block for random probability measures (James [18]) and has been widely investigated (Lijoi *et al.* [34]). Earlier studies on this process can be found in modeling survival functions (Hougaard [15]) and in spatial modeling (Wolpert and Ickstadt [45]; Brix [4]). The intensity of the generalized gamma process is given by

$$\nu(dx, ds) = \eta(dx) \times \rho_x(ds) = \eta(dx) \times \frac{\theta}{\Gamma(1 - \alpha)} s^{-1-\alpha} e^{-b(x)s} ds,$$

$$0 \leq \alpha < 1, \theta > 0, b(x) \geq 0.$$

This process is an important class since it contains the following well known processes:

1. Taking $0 \leq \alpha < 1, \theta > 0, b(x) = 0$ yields the intensity of positive α stable process,
2. Taking $\alpha = 0, \theta > 0, b(x) = b \geq 0$ yields the intensity of gamma process, where b is a known constant,
3. Taking $\alpha = 0, \theta > 0, b(x) \geq 0$ yields the intensity of extended gamma process and
4. Taking $\alpha = 1/2, \theta > 0, b(x) = b \geq 0$ yields the intensity of inverse Gaussian process, where b is a known constant.

We now consider the normalized random measure derived from the polynomially and exponentially tilted law of the generalized gamma process. An application of Theorem 3.2 and Proposition 3.1 leads to the conditional GBMP urn formula

$$\begin{aligned} & \mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} \\ &= \left(\frac{\theta}{(\gamma + \tilde{U}_n + b(x))^{1-\alpha}} \right. \\ & \left. / \left(\int_{\mathbb{X}} \frac{\theta}{(\gamma + \tilde{U}_n + b(x))^{1-\alpha}} \eta(dx) + \sum_{k'=1}^{N(\mathbf{p}_n)} \frac{n_{k'} - \alpha}{\gamma + \tilde{U}_n + b(Y_{k'})} \right) \right) \eta(dx) \end{aligned} \tag{4.1}$$

$$\begin{aligned}
 & + \sum_{k=1}^{S(\mathbf{p}_n)} \left(\frac{n_k - \alpha}{\gamma + \tilde{U}_n + b(Y_k)} \right. \\
 & \left. / \left(\int_{\mathbb{X}} \frac{\theta}{(\gamma + \tilde{U}_n + b(x))^{1-\alpha}} \eta(dx) + \sum_{k'=1}^{S(\mathbf{p}_n)} \frac{n_{k'} - \alpha}{\gamma + \tilde{U}_n + b(Y_{k'})} \right) \right) \delta_{Y_k}(dx),
 \end{aligned}$$

and the posterior distribution of \tilde{U}_n is given by

$$f_{\tilde{U}_n | \mathbf{X}_n}(u) \propto \begin{cases} \left[\int_{\mathbb{X}} \frac{\theta}{(\gamma + u + b(x))^{1-\alpha}} \eta(dx) + \sum_{k'=1}^{S(\mathbf{p}_n)} \frac{n_{k'} - \alpha}{\gamma + u + b(Y_{k'})} \right] \\ \times \frac{e^{-(\theta/\alpha) \int_{\mathbb{X}} (\gamma + u + b(x))^\alpha \eta(dx)} u^{n+q}}{\prod_{k=1}^{S(\mathbf{p}_n)} (\gamma + u + b(Y_k))^{n_k - \alpha}}, & 0 < \alpha < 1, \\ \left[\int_{\mathbb{X}} \frac{\theta}{\gamma + u + b(x)} \eta(dx) + \sum_{k'=1}^{S(\mathbf{p}_n)} \frac{n_{k'}}{\gamma + u + b(Y_{k'})} \right] \\ \times \frac{e^{-\theta \int_{\mathbb{X}} \ln(\gamma + u + b(x)) \eta(dx)} u^{n+q}}{\prod_{k=1}^{S(\mathbf{p}_n)} (\gamma + u + b(Y_k))^{n_k}}, & \alpha = 0. \end{cases}$$

Notice that when $b(x) = b > 0$ and $\alpha = 0$, then the condition (2.5) is reduced to $\theta > q \geq 0$. In fact, when $b(x) = b > 0$ and $0 < \alpha < 1$, then θ is not required to be greater than q and the condition (2.5) holds with $q \geq 0$ and $\theta > 0$. With a general $b(x)$, the condition (2.5) is required to be examined. By inspection of (4.1), one could find that when $\alpha = 0$ and $b(x) = b$, X_{n+1} is no longer dependent on \tilde{U}_n given the past $\{X_1, \dots, X_n\}$, that is $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} = \mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\}$. This fact is also emphasized in Remark 2 of James *et al.* [22], page 86. This setting with $q = 0$ is corresponding to the normalized gamma process or the Dirichlet process and the GBMP urn formula is given by

$$\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} = \frac{\theta}{\theta + n} \eta(dx) + \sum_{k=1}^{S(\mathbf{p}_n)} \frac{n_k}{\theta + n} \delta_{Y_k}(dx). \tag{4.2}$$

This (4.2) is the BMP urn formula (see also Blackwell and MacQueen [3]).

Normalized beta-gamma process

The law of the beta-gamma process can be derived from the polynomially tilting the law of the gamma process (James [20]). The beta-gamma process was first introduced for the proof of the Markov–Krein identity of the Dirichlet process mean functionals and since then it has become a useful analytical tool for studying the Dirichlet process (see James [20] and James *et al.* [22]). Here we mention that Cifarelli and Regazzini [5–7] were the first works in which this identity was explicitly demonstrated in relation to the research of the law of the Dirichlet process mean functionals. An important fact in James [20] states that a Dirichlet process can be expressed

as a normalized beta-gamma process. This becomes an interesting alternative expression of the Dirichlet process that is usually expressed as the normalized gamma process. This expression is given by taking $q > 0, \gamma = 0, \alpha = 0, \theta > q, b(x) = b \geq 0$. The condition $\theta > q$ (see also James [20], equation 5, page 649) is equivalent to the condition (2.5). With $\alpha = 0$ and $b(x) = b$, as in the construction through the normalized gamma process, the variable X_{n+1} is not dependent on \tilde{U}_n nor U_n given the past $\{X_1, \dots, X_n\}$.

Poisson–Dirichlet process

The Poisson–Dirichlet process is a common and well known process used in both statistical and probabilistic modeling. This is also called Pitman–Yor process which is coined by Ishwaran and James [16]. This process has been shown to be useful in a variety of interesting applications in combinatorics (Arratia *et al.* [2]), population genetics (Griffiths and Lessard [14]) and Bayesian statistics (Ishwaran and James [16,17]). This process was first introduced by Kingman [27] and Pitman and Yor [43] provided a detailed study of its properties. We consider the Poisson–Dirichlet process with parameters (α, q) that is equivalent to take $\theta = 1, \gamma = 0, b(x) = 0, q > 0$, and $0 < \alpha < 1$. Special cases include the Dirichlet process and the normalized stable process with parameters $(0, q)$ and $(\alpha, 0)$ respectively. These two processes could be seen as an two parameter extension of the Dirichlet process. Even though there is an explicit expression of the unconditional GBMP urn formula (Pitman [41,42]; Ishwaran and James [16,17]), it is still worth examining the augmented version for the Poisson–Dirichlet process (α, q) process. The conditional GBMP urn formula is given by

$$\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} = \frac{\tilde{U}_n^\alpha}{\tilde{U}_n^\alpha + n - \alpha \mathfrak{N}(\mathbf{p}_n)} \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{n_k - \alpha}{\tilde{U}_n^\alpha + n - \alpha \mathfrak{N}(\mathbf{p}_n)} \delta_{Y_k}(dx),$$

where

$$f_{\tilde{U}_n | \mathbf{X}_n}(u) = \frac{q + \alpha \mathfrak{N}(\mathbf{p}_n)}{n + q} \frac{e^{-(1/\alpha)u^\alpha} u^{q + \alpha \mathfrak{N}(\mathbf{p}_n) + \alpha - 1}}{\alpha^{q/\alpha + \mathfrak{N}(\mathbf{p}_n)} \Gamma(q/\alpha + \mathfrak{N}(\mathbf{p}_n) + 1)} + \frac{n - \alpha \mathfrak{N}(\mathbf{p}_n)}{n + q} \frac{e^{-(1/\alpha)u^\alpha} u^{q + \alpha \mathfrak{N}(\mathbf{p}_n) - 1}}{\alpha^{q/\alpha + \mathfrak{N}(\mathbf{p}_n) - 1} \Gamma(q/\alpha + \mathfrak{N}(\mathbf{p}_n))}. \tag{4.3}$$

This is equivalent to $\tilde{U}_n \stackrel{d}{=} G_1^{1/\alpha}$ with probability $(q + \alpha \mathfrak{N}(\mathbf{p}_n))/(n + q)$ and $\tilde{U}_n \stackrel{d}{=} G_2^{1/\alpha}$ with probability $(n - \alpha \mathfrak{N}(\mathbf{p}_n))/(n + q)$ where G_1 is a Gamma($q/\alpha + \mathfrak{N}(\mathbf{p}_n) + 1, 1/\alpha$) random variable and G_2 is a Gamma($q/\alpha + \mathfrak{N}(\mathbf{p}_n), 1/\alpha$) random variable.

Normalized generalized gamma process

The normalized generalized gamma process with $\theta > 0, \gamma = 0, b(x) = b > 0, q = 0$, and $0 < \alpha < 1$ is considered in James [18] and Lijoi *et al.* [32,34]. Specifically, the conditional GBMP urn

formula is given by

$$\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} = \frac{\theta(\tilde{U}_n + b)^\alpha}{\theta(\tilde{U}_n + b)^\alpha + n - \alpha \mathfrak{N}(\mathbf{p}_n)} \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{n_k - \alpha}{\theta(\tilde{U}_n + b)^\alpha + n - \alpha \mathfrak{N}(\mathbf{p}_n)} \delta_{Y_k}(dx),$$

where the density of the augmented variable \tilde{U}_n is given by

$$f_{\tilde{U}_n | \mathbf{X}_n}(u) \propto \frac{[\theta(u + b)^\alpha + n - \alpha \mathfrak{N}(\mathbf{p}_n)] e^{-(\theta/\alpha)(u+b)^\alpha} u^n}{(u + b)^{n - \alpha \mathfrak{N}(\mathbf{p}_n) + 1}}. \tag{4.4}$$

A little effort might be required for sampling \tilde{U}_n from (4.4). One could follow Devroye [9], Section II.3.3, page 47, to derive a suitable rejection procedure. Here we give a simple illustration. A sample of V_n could be drawn from

$$\frac{\alpha \mathfrak{N}(\mathbf{p}_n)}{n} \frac{\theta^{\mathfrak{N}(\mathbf{p}_n) + 1} e^{-(\theta/\alpha)v^\alpha} v^{\alpha \mathfrak{N}(\mathbf{p}_n) + \alpha - 1}}{\alpha^{\mathfrak{N}(\mathbf{p}_n)} \Gamma(\mathfrak{N}(\mathbf{p}_n) + 1)} + \frac{n - \alpha \mathfrak{N}(\mathbf{p}_n)}{n} \frac{\theta^{\mathfrak{N}(\mathbf{p}_n)} e^{-(\theta/\alpha)v^\alpha} v^{\alpha \mathfrak{N}(\mathbf{p}_n) - 1}}{\alpha^{\mathfrak{N}(\mathbf{p}_n) - 1} \Gamma(\mathfrak{N}(\mathbf{p}_n))},$$

and if $\zeta < \psi(V_n)$, then $\tilde{U}_n = V_n$, otherwise sample V_n again until $\zeta < \psi(V_n)$ where ζ is an uniform random variable which is independent of V_n and $\psi(v) = e^{-(\theta/\alpha)[(v+b)^\alpha - v^\alpha]} \times (\frac{v}{v+b})^{n - \alpha \mathfrak{N}(\mathbf{p}_n) + 1} \times \frac{\theta(v+b)^\alpha + (n - \alpha \mathfrak{N}(\mathbf{p}_n))}{\theta v^\alpha + (n - \alpha \mathfrak{N}(\mathbf{p}_n))}$. Notice that \tilde{U}_n given \mathbf{X}_n (4.4) and V_n are identical in distribution when $b = 0$. The random variable V_n can be described as: $V_n \stackrel{d}{=} G_1^{1/\alpha}$ with probability $\alpha \mathfrak{N}(\mathbf{p}_n)/n$ and $V_n \stackrel{d}{=} G_2^{1/\alpha}$ with probability $(n - \alpha \mathfrak{N}(\mathbf{p}_n))/n$ where G_1 is a Gamma($\mathfrak{N}(\mathbf{p}_n) + 1, \theta/\alpha$) random variable and G_2 is a Gamma($\mathfrak{N}(\mathbf{p}_n), \theta/\alpha$) random variable.

4.2. Generalized Dirichlet process

Regazzini *et al.* [44] introduce the generalized Dirichlet process as an example for determining the mean of normalized random measures with independent increments. Apart from studying probabilistic properties of the generalized Dirichlet process, its use in Bayesian non-parametric statistics is developed in Lijoi *et al.* [31]. We state the intensity of the generalized Dirichlet process as

$$v(ds, dx) = \eta(dx) \times \rho_x(ds) = \eta(dx) \times \theta \frac{1 - e^{-cs}}{1 - e^{-s}} s^{-1} e^{-s} ds,$$

and introduce the difference of two Hurwitz Zeta functions as,

$$\varphi_{n_k}(\gamma + u, c) = \sum_{\ell=0}^{\infty} \left[\frac{1}{(\gamma + u + 1 + \ell)^{n_k}} - \frac{1}{(\gamma + u + c + 1 + \ell)^{n_k}} \right].$$

When c is a positive integer (which is considered by Regazzini *et al.* [44] and Lijoi *et al.* [31]), the function can be simplified to a finite sum,

$$\varphi_{n_k}(\gamma + u, c) = \sum_{\ell=0}^{c-1} \frac{1}{(\gamma + u + 1 + \ell)^{n_k}}, \quad c = 1, 2, \dots$$

The conditional GBMP urn is given by

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} &= \left(\theta \varphi_1(\gamma + \tilde{U}_n, c) \right. \\ &\quad \left. / \left(\theta \varphi_1(\gamma + \tilde{U}_n, c) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} n_{k'} \frac{\varphi_{n_{k'}+1}(\gamma + \tilde{U}_n, c)}{\varphi_{n_{k'}}(\gamma + \tilde{U}_n, c)} \right) \right) \eta(dx) \\ &\quad + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \left(n_k \frac{\varphi_{n_k+1}(\gamma + \tilde{U}_n, c)}{\varphi_{n_k}(\gamma + \tilde{U}_n, c)} \right. \\ &\quad \left. / \left(\theta \varphi_1(\gamma + \tilde{U}_n, c) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} n_{k'} \frac{\varphi_{n_{k'}+1}(\gamma + \tilde{U}_n, c)}{\varphi_{n_{k'}}(\gamma + \tilde{U}_n, c)} \right) \right) \delta_{Y_k}(dx), \end{aligned}$$

where the density of the augmented variable is given by

$$\begin{aligned} f_{\tilde{U}_n | \mathbf{X}_n}(u) &\propto \chi(u) \\ &= \left[\theta \varphi_1(\gamma + u, c) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} n_{k'} \frac{\varphi_{n_{k'}+1}(\gamma + u, c)}{\varphi_{n_{k'}}(\gamma + u, c)} \right] \\ &\quad \times \left[\frac{\Gamma(\gamma + u + 1)}{\Gamma(\gamma + u + c + 1)} \right]^\theta u^{n+q} \times \theta^{\mathfrak{N}(\mathbf{p}_n)} \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \varphi_{n_k}(\gamma + u, c). \end{aligned}$$

Again, the condition (2.5) stated in the construction that $E[h(\mu(\mathbb{X}))] < \infty$ is equivalent to $\theta > q$. In particular, when $c = 1$ this is corresponding to the Dirichlet process, $\varphi_1(\gamma + u, 1) = \varphi_{n_k+1}(\gamma + u, 1)/\varphi_{n_k}(\gamma + u, 1)$ and it can be shown that X_{n+1} given $\{X_1, \dots, X_n\}$ is not dependent on U_n . Similar to the normalized generalized gamma process, the rejection method can also be proposed due to Devroye [9], Section II.3.3, page 47. That is, sample W_n from Beta($n + q + 1, \theta - q$), and if $\zeta < \psi((\gamma + 1)W_n/(1 - W_n))$, then $\tilde{U}_n = (\gamma + 1)W_n/(1 - W_n)$, otherwise sample W_n again until $\zeta < \psi((\gamma + 1)W_n/(1 - W_n))$ where ζ is a uniform random variable which is independent of W_n and $\psi(v) = \chi(v) \times (\gamma + v + 1)^{\theta+n+1}/(c^{\mathfrak{N}(\mathbf{p}_n)}(c\theta + n))$.

5. Simulation study on approximating distribution of partition size via the augmented Blackwell–MacQueen Pólya urn formula

We conduct a simulation study on approximating the posterior probabilities of partition sizes, $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_n) = i\}$ for $i = 1, \dots, n$, using the conditional GBMP urn formula discussed in last section. We consider two popular random probability measures, the Poisson–Dirichlet process and the normalized generalized gamma process. For the Poisson–Dirichlet process, we simulate data according to $\alpha = 0.5, \theta = 1, q = 1, \gamma = 0, b(x) = 0$. For the normalized generalised gamma process, we simulate data according to $\alpha = 0.5, \theta = 1, q = 0, \gamma = 0, b(x) = 1$. In both cases, we set $n = 50$. We examine two exact sequential sampling schemes and two MCMC schemes. Specifically:

- A.1. Sample X_i sequentially for $i = 1, \dots, n$ according to the unconditional GBMP urn formula, $\mathbb{P}\{X_{i+1} \in dx | \mathbf{X}_i\}$ for $i = 1, \dots, n - 1$.
- A.2. Sample $\{X_i, \tilde{U}_i\}$ sequentially for $i = 1, \dots, n$ according to the conditional GBMP urn formula, $\mathbb{P}\{X_{i+1} \in dx | \mathbf{X}_i, \tilde{U}_i\}$ and $f_{\tilde{U}_i | \mathbf{X}_i}(u)$ for $i = 1, \dots, n - 1$.
- A.3. Re-sample X_i iteratively for $i = 1, \dots, n$ according to the unconditional GBMP urn formula, $\mathbb{P}\{X_i \in dx | \mathbf{X}_n \setminus \{X_i\}\}$ for $i = 1, \dots, n$.
- A.4. Re-sample $\{X_i, U_n\}$ iteratively for $i = 1, \dots, n$ according to the conditional GBMP urn formula, $\mathbb{P}\{X_i \in dx | \mathbf{X}_n \setminus \{X_i\}, U_n\}$ and $f_{U_n | \mathbf{X}_n}(u)$ for $i = 1, \dots, n$.

Algorithms A.1 and A.2 are exact and are identical in distribution. Algorithm A.1 has been frequently used in the literature and for the normalized generalized gamma process requires the evaluation of the complicated functions. Algorithm A.2 is the conditional GBMP urn formula derived from the tilted measure proposed in this article. This method is straightforward to implement without much complication in evaluations. We also include two MCMC Gibbs sampling algorithms described in A.3 and A.4 for comparison. These are not exact sampling algorithms and an initial sampling period is necessary to converge to the stationary distribution. The stationary distribution itself is identical to that of A.1 and A.2.

In each replication, we sample 10,000 independent samples from A.1 and A.2 to approximate $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_n) = i\}$ for $i = 1, \dots, n$. Starting with a partition $\mathbf{p}_n = \{\{1\}, \dots, \{n\}\}$ with all singleton clusters, we draw 20,000 samples from algorithms A.3 and A.4. We ignore the first 10,000 warmup samples and use the last 10,000 samples to approximate $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_n) = i\}$ for $i = 1, \dots, n$. So, each algorithm produces 10,000 approximates of probabilities of partition sizes, $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_n) = 1\}, \dots, \mathbb{P}\{\mathfrak{N}(\mathbf{p}_n) = n\}$. To summarize the results, Figures 1 and 2 show the range and 95% confidence levels of 10,000 approximates of probabilities of partition sizes for the Poisson–Dirichlet process and the normalized generalized gamma process respectively for algorithms A.1–A.4. Similarly, Tables 1 and 2 shows the true probabilities and the means and standard errors of the approximates given by algorithms A.1–A.4.

Figure 1 and Table 1 show that for the Poisson–Dirichlet process algorithms A.1 and A.2 result in samples from identical distributions, as the theory would suggest. The MCMC algorithms, A.3 and A.4 also produce similar results to A.1 and A.2 except for the standard errors in Table 1. The standard errors indicate the variability of the MCMC generated samples is generally greater than

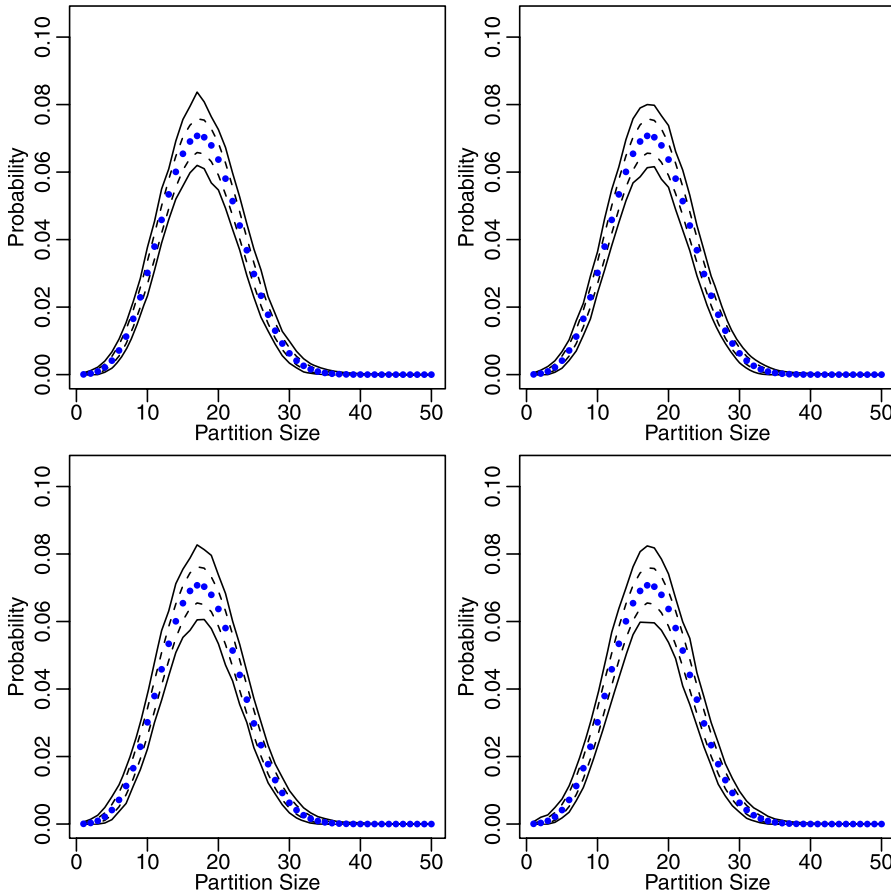


Figure 1. Approximate probabilities of partition sizes of the Poisson Dirichlet process ($\alpha = 0.5, \theta = 1, q = 1, \gamma = 0, b(x) = 0$), $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_{50}) = i\}$ for $i = 1, \dots, 50$. Top Left: A.1 algorithm; Top Right: A.2 algorithm; Bottom Left: A.3 algorithm; Bottom Right: A.4 algorithm. The solid bound lines indicate the range of all 10,000 approximates of the probabilities; The dash bound lines indicate the 95% confidence level (2.5% and 97.5% quantiles) of all 10,000 approximates of the probabilities. The dots indicate the true probabilities.

the exact sequential sampling algorithms, as we would expect. The results for the normalized generalized gamma process shown in Figure 2 and Table 2 point to a similar story as in the Poisson–Dirichlet process.

Finally, we note that it is not necessary to sample \mathbf{X}_n to conduct the simulation. It could be done by simply simulating partitions \mathbf{p}_n directly instead. It is possible to integrate out all $\mathbf{Y}_{\mathfrak{N}(\mathbf{p}_n)}$ from the conditional GBMP urn formula and obtain the weights for partition sampling using the Chinese restaurant process (see Aldous [1], Lo *et al.* [38] and Pitman [42]).

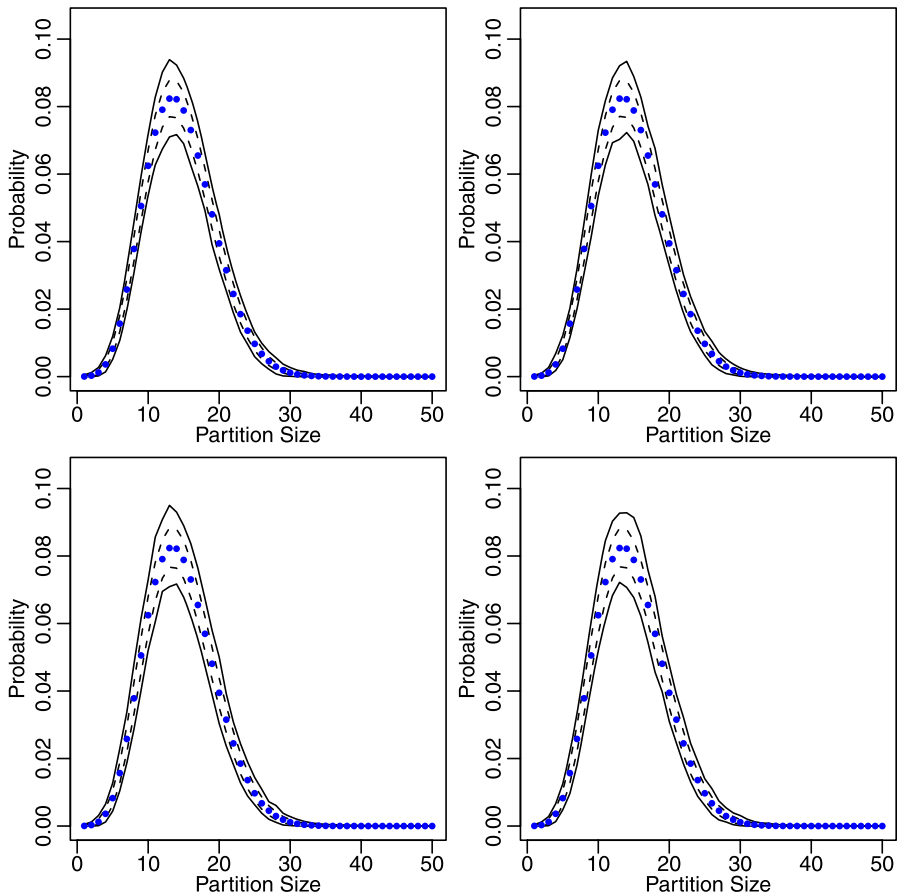


Figure 2. Approximate probabilities of partition sizes of the normalized generalized Gamma process ($\alpha = 0.5$, $\theta = 1$, $q = 0$, $\gamma = 0$, $b(x) = 1$), $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_{50}) = i\}$ for $i = 1, \dots, 50$. Top Left: A.1 algorithm; Top Right: A.2 algorithm; Bottom Left: A.3 algorithm; Bottom Right: A.4 algorithm. The solid bound lines indicate the range of all 10,000 approximates of the probabilities; The dash bound lines indicate the 95% confidence level (2.5% and 97.5% quantiles) of all 10,000 approximates of the probabilities. The dots indicate the true probabilities.

6. Conclusion and further research

This article has introduced a class of random probability measures based on polynomially and exponentially tilting. We have provided a complete Bayesian analysis of this class of measures with details on the prior and posterior laws and shown that the class is structurally conjugate. We described a conditional Blackwell–MacQueen Pólya urn sampling scheme that simplifies the computational requirements to implement such sampling schemes. The new sampling scheme yields similar answers to more complicated schemes described in the literature.

Table 1. Approximate probabilities of partition sizes of the Poisson Dirichlet process ($\alpha = 0.5, \theta = 1, q = 1, \gamma = 0, b(x) = 0$), $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_{50}) = i\}$ for $i = 1, \dots, 50$. *First Column:* Partition sizes; *Second Column:* True probability of partition sizes; *Third Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.1; *Fourth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.1; *Fifth Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.2; *Sixth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.2; *Seventh Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.3; *Eighth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.3; *Ninth Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.4; *Tenth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.4

$\mathbb{P}\{\mathfrak{N}(\mathbf{p}_n) = i\}$		A.1 algorithm		A.2 algorithm		A.3 algorithm		A.4 algorithm	
i	True probability	Mean of the approximates	SE of the approximates	Mean of the approximates	SE of the approximates	Mean of the approximates	SE of the approximates	Mean of the approximates	SE of the approximates
1	0.000063	0.000062	0.000080	0.000063	0.000079	0.000063	0.000088	0.000064	0.000088
2	0.000315	0.000316	0.000179	0.000318	0.000179	0.000313	0.000207	0.000311	0.000207
3	0.000936	0.000933	0.000306	0.000934	0.000308	0.000934	0.000369	0.000936	0.000371
4	0.002139	0.002138	0.000462	0.002133	0.000460	0.002138	0.000567	0.002141	0.000575
5	0.004142	0.004146	0.000639	0.004140	0.000641	0.004147	0.000806	0.004135	0.000807
6	0.007139	0.007129	0.000837	0.007143	0.000836	0.007139	0.001071	0.007131	0.001068
7	0.011258	0.011256	0.001060	0.011242	0.001066	0.011257	0.001351	0.011258	0.001333
8	0.016537	0.016519	0.001274	0.016563	0.001277	0.016536	0.001613	0.016526	0.001610
9	0.022898	0.022916	0.001510	0.022901	0.001502	0.022883	0.001868	0.022890	0.001849
10	0.030135	0.030133	0.001705	0.030123	0.001716	0.030089	0.002075	0.030119	0.002101
11	0.037923	0.037944	0.001923	0.037937	0.001905	0.037893	0.002248	0.037910	0.002311
12	0.045836	0.045829	0.002081	0.045820	0.002121	0.045812	0.002430	0.045901	0.002421
13	0.053388	0.053358	0.002248	0.053434	0.002243	0.053396	0.002515	0.053364	0.002539
14	0.060073	0.060076	0.002407	0.060042	0.002371	0.060038	0.002572	0.060061	0.002600
15	0.065424	0.065423	0.002486	0.065396	0.002497	0.065402	0.002638	0.065415	0.002683
16	0.069059	0.069067	0.002527	0.069071	0.002554	0.069094	0.002677	0.069051	0.002689
17	0.070723	0.070738	0.002575	0.070720	0.002555	0.070754	0.002673	0.070746	0.002684
18	0.070317	0.070311	0.002544	0.070328	0.002558	0.070302	0.002729	0.070290	0.002691
19	0.067906	0.067983	0.002499	0.067911	0.002480	0.067937	0.002692	0.067889	0.002688
20	0.063706	0.063723	0.002443	0.063674	0.002479	0.063756	0.002640	0.063700	0.002636
21	0.058061	0.058033	0.002325	0.058056	0.002349	0.058063	0.002585	0.058066	0.002583

Table 2. Approximate of probabilities of partition sizes of the normalized generalized Gamma process ($\alpha = 0.5, \theta = 1, q = 0, \gamma = 0, b(x) = 1$), $\mathbb{P}\{\mathfrak{N}(\mathbf{p}_{50}) = i\}$ for $i = 1, \dots, 50$. *First Column:* Partition sizes; *Second Column:* True probability of partition sizes; *Third Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.1; *Fourth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.1; *Fifth Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.2; *Sixth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.2; *Seventh Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.3; *Eighth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.3; *Ninth Column:* Mean of 10,000 approximates of the probabilities according to algorithm A.4; *Tenth Column:* Standard error of 10,000 approximates of the probabilities according to algorithm A.4

i	$\mathbb{P}\{\mathfrak{N}(\mathbf{p}_n) = i\}$ True probability	A.1 algorithm		A.2 algorithm		A.3 algorithm		A.4 algorithm	
		Mean of the approximates	SE of the approximates	Mean of the approximates	SE of the approximates	Mean of the approximates	SE of the approximates	Mean of the approximates	SE of the approximates
1	0.000035	0.000036	0.000059	0.000035	0.000060	0.000035	0.000061	0.000035	0.000061
2	0.000291	0.000294	0.000170	0.000290	0.000170	0.000288	0.000178	0.000293	0.000183
3	0.001234	0.001227	0.000353	0.001238	0.000356	0.001229	0.000389	0.001241	0.000396
4	0.003621	0.003621	0.000599	0.003624	0.000604	0.003620	0.000701	0.003622	0.000703
5	0.008265	0.008274	0.000916	0.008266	0.000909	0.008249	0.001101	0.008262	0.001083
6	0.015684	0.015709	0.001237	0.015683	0.001250	0.015647	0.001554	0.015666	0.001551
7	0.025793	0.025807	0.001598	0.025753	0.001594	0.025786	0.001989	0.025825	0.001997
8	0.037832	0.037827	0.001887	0.037807	0.001896	0.037806	0.002365	0.037843	0.002367
9	0.050538	0.050540	0.002207	0.050498	0.002233	0.050500	0.002645	0.050583	0.002657
10	0.062458	0.062450	0.002393	0.062502	0.002396	0.062445	0.002807	0.062497	0.002823
11	0.072282	0.072260	0.002591	0.072270	0.002557	0.072238	0.002921	0.072246	0.002880
12	0.079070	0.079074	0.002674	0.079123	0.002698	0.079058	0.002932	0.079112	0.002914
13	0.082360	0.082371	0.002745	0.082353	0.002724	0.082412	0.002945	0.082404	0.002924
14	0.082159	0.082126	0.002758	0.082176	0.002764	0.082174	0.002894	0.082143	0.002915
15	0.078842	0.078821	0.002708	0.078836	0.002675	0.078878	0.002847	0.078786	0.002846
16	0.073037	0.072990	0.002618	0.073001	0.002614	0.073106	0.002783	0.073063	0.002804
17	0.065486	0.065510	0.002469	0.065512	0.002465	0.065489	0.002706	0.065455	0.002707
18	0.056942	0.056976	0.002343	0.056984	0.002316	0.056960	0.002585	0.056956	0.002560
19	0.048085	0.048085	0.002165	0.048092	0.002160	0.048113	0.002429	0.048047	0.002398
20	0.039472	0.039478	0.001975	0.039484	0.001980	0.039473	0.002225	0.039428	0.002223
21	0.031516	0.031532	0.001743	0.031496	0.001741	0.031529	0.002037	0.031498	0.002037

We also note a general tilting treatment could be considered for any measurable function h on \mathbb{R}^+ evaluated at the total mass of the corresponding CRM in (2.6). This general class of random probability measures with homogeneous intensity, $\nu(dx, ds) = \eta(dx)\rho(ds)$, can be shown to be the Poisson–Kingman process (Pitman [41,42]) by showing that the normalized random probability measure (2.6) has a Poisson–Kingman partition (ρ, γ) , where γ denotes the tilted density of total mass such that $\gamma = f_{\tilde{T}}$ (2.8) and ρ represents the σ -finite non-atomic intensity measure $\rho(ds)$. Then, the conditional partition distribution is given by

$$\mathbb{P}(\mathbf{p}_n | t) = \int_{(\mathbb{R}^+)^{\mathfrak{N}(\mathbf{p}_n)}} \mathbb{I}_{\{t - \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} s_k > 0\}} \frac{f_{\tilde{T}}(t - \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} s_k)}{t^n f_{\tilde{T}}(t)} \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} s_k^{n_k} \rho(ds_k). \tag{6.1}$$

The proof is given in Appendix A.5; See also Pitman [41], Lemma 5, page 8, for the related joint distribution of partitions and jumps. Investigation of the properties of this general class of random probably measures and even special cases of (2.6) is interesting for further research.

Finally, we note that the applications of Bayesian non-parametric mixture models in Bayesian statistics is steadily increasing (see Lo [36]; James [18]; Ishwaran and James [16,17]; James *et al.* [22]; Lijoi *et al.* [31,32,34]). In particular, time series model mixing over random probability measures has been considered recently in Griffin and Steel [13], Lau and So [29,30] and Lau and Cripps [28]. Often in mixture models over the normalized tilted CRM, it is necessary to consider a collection of latent variables, which is a sample of the normalized tilted CRM but these latent variables are not observed directly. In fact, sampling latent variables is essential for approximating estimates of parameters of interest that are functions of the latent variables, \mathbf{X}_n because of the high cardinality of the posterior distribution due to combinatorial property of the latent variables. As a result, sampling schemes for \mathbf{X}_n are required for estimation based on the conditional Blackwell–MacQueen Pólya urn formula and the distributions of \mathbf{X}_n and partitions \mathbf{p}_n . In this article, we have provided the marginal distributions of \mathbf{X}_n and \mathbf{p}_n , both conditional and unconditional on U_n , that are essential elements in implementing mixture models over the normalized tilted CRM.

Appendix

A.1. Proof of Theorem 3.1

Consider the joint distribution of $(\tilde{\mu}_t, \mathbf{X}_n)$, which is given by (3.4) and (3.5), that is

$$\mathcal{P}_{\tilde{\mu}_t, \mathbf{X}_n}(A, B_1, \dots, B_n) = \frac{1}{E[h'(\tilde{\mu}(\mathbb{X}))\tilde{\mu}(\mathbb{X})^{-q}]} \int_A \frac{h'(\mu(\mathbb{X}))}{\mu(\mathbb{X})^{n+q}} \left(\prod_{i=1}^n \mu(B_i) \right) \mathcal{P}_{\tilde{\mu}}(d\mu). \tag{A.1}$$

In (A.1), the CRM $\tilde{\mu}(\cdot)$ is now replaced by the linear functional of the Poisson process $\int_{\mathbb{R}^+} sN(\cdot, ds)$ on \mathbb{X} (2.4). Following by writing the right hand side of (A.1) without the integrals, that is given by

$$\frac{h'(\int_{\mathbb{X} \times \mathbb{R}^+} sN(dx, ds))}{(\int_{\mathbb{X} \times \mathbb{R}^+} sN(dx, ds))^{n+q}} \left(\prod_{i=1}^n s_i N(dx_i, ds_i) \right) \mathcal{P}_N(dN), \tag{A.2}$$

where this Poisson process N has distribution denoted by \mathcal{P}_N , has the intensity measure ν same as that of $\tilde{\mu}$ (2.4), and belong to the set in $\mathcal{B}(\mathcal{M}_{\mathbb{X} \times \mathbb{R}^+})$ corresponding to the set A that $\mu \in A \in \mathcal{B}(\mathcal{M}_{\mathbb{X}})$. Here (x_i, s_i) for $i = 1, \dots, n$ represent the points generated from the Poisson process N . For $k = 1, \dots, \aleph(\mathbf{p}_n)$, (Y_k, J_k) denotes the distinct points of the (x_i, s_i) 's and $\aleph(\mathbf{p}_n)$ denotes the number of the distinct points. In addition, we take $\{J_1, \dots, J_{\aleph(\mathbf{p}_n)}\}$ to be the augmented variables. We then apply the Fubini theorem following from an application of Lemma 2.2 of James [18], page 8 (see also James [18] for some detail discussions), to yield the joint distribution of N, \mathbf{p}_n and (Y_k, J_k) for $k = 1, \dots, \aleph(\mathbf{p}_n)$, which is given by (A.2) without the proportional constant,

$$\frac{h'(\int_{\mathbb{X} \times \mathbb{R}^+} sN(\mathbf{X}_n)(dx, ds))}{(\int_{\mathbb{X} \times \mathbb{R}^+} sN(\mathbf{X}_n)(dx, ds))^{n+q}} \mathcal{P}_N(dN) \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k) s_k^{n_k} \rho_{Y_k}(ds_k), \tag{A.3}$$

where $N(\mathbf{X}_n) \stackrel{d}{=} N + \sum_{k=1}^{\aleph(\mathbf{p}_n)} \delta_{(Y_k, J_k(\mathbf{X}_n))}$. Here the distribution of $N(\mathbf{X}_n)$ is identical to the distribution of the sum over a Poisson process N and the fixed points of discontinuity at $(Y_k, J_k(\mathbf{X}_n))$. The Poisson process N has the intensity measure ν and be independent of $(Y_k, J_k(\mathbf{X}_n))$ for $k = 1, \dots, \aleph(\mathbf{p}_n)$. The pairs $(Y_k, J_k(\mathbf{X}_n))$ for $k = 1, \dots, \aleph(\mathbf{p}_n)$ and \mathbf{p}_n has the joint distribution $\prod_{k=1}^{\aleph(\mathbf{p}_n)} [\eta(dY_k) \times \mathbb{P}(J_k(\mathbf{X}_n) \in ds_k) \times \tau_{n_k, Y_k}(0)]$ where $\mathbb{P}(J_k(\mathbf{X}_n) \in ds_k) = s_k^{n_k} \rho_{Y_k}(ds_k) / \int_{\mathbb{R}^+} s_k^{n_k} \rho_{Y_k}(ds_k)$ and $\tau_{n_k, Y_k}(0)$ defined in (3.17). The distribution of $N(\mathbf{X}_n)$ now is denoted by $\mathcal{P}_N^{(\mathbf{X}_n)}$, so (A.3) could be reduced to

$$\frac{h'(\int_{\mathbb{X} \times \mathbb{R}^+} sN(dx, ds))}{(\int_{\mathbb{X} \times \mathbb{R}^+} sN(dx, ds))^{n+q}} \mathcal{P}_N^{(\mathbf{X}_n)}(dN). \tag{A.4}$$

The Poisson linear functional appears in (A.4), $\int_{\mathbb{R}^+} sN(\mathbf{X}_n)(dx, ds)$ ($=: \tilde{\mu}(\mathbf{X}_n)(dx)$), is a completely random measure according to (2.1) (see also Daley and Vere-Jones [8], Theorem 10.1.III, page 79), such that

$$\tilde{\mu}(\mathbf{X}_n)(dx) \stackrel{d}{=} \int_{\mathbb{R}^+} sN(dx, ds) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} J_k(\mathbf{X}_n) \delta_{Y_k}(dx) \stackrel{d}{=} \tilde{\mu}(dx) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} J_k(\mathbf{X}_n) \delta_{Y_k}(dx), \tag{A.5}$$

where $\tilde{\mu}$ denote a CRM (2.4) derived from the Poisson process N . So, the total mass $\tilde{\mu}(\mathbf{X}_n)(\mathbb{X})$ in (A.4) is given by

$$\tilde{\mu}(\mathbf{X}_n)(\mathbb{X}) \stackrel{d}{=} \tilde{\mu}(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} J_k(\mathbf{X}_n). \tag{A.6}$$

This immediately reveals the distribution of the tilted completely random measure, $\tilde{\mu}_t^{(\mathbf{X}_n)}$ (Definition 2.1), that is $\mathcal{P}_{\tilde{\mu}_t^{(\mathbf{X}_n)}}(d\mu) \propto \frac{h'(\mu(\mathbf{X}_n)(\mathbb{X}))}{(\mu(\mathbf{X}_n)(\mathbb{X}))^{n+q}} \mathcal{P}_{\tilde{\mu}}^{(\mathbf{X}_n)}(d\mu)$. Lastly, the distributional identity derived from the a sample $\tilde{\mu}_t^{(\mathbf{X}_n)}$ from $\mathcal{P}_{\tilde{\mu}_t^{(\mathbf{X}_n)}}^{(\mathbf{X}_n)}$, $G(\mathbf{X}_n)(\cdot) \stackrel{d}{=} \tilde{\mu}_t^{(\mathbf{X}_n)}(\cdot) / \tilde{\mu}_t^{(\mathbf{X}_n)}(\mathbb{X})$ on \mathbb{X} is obtained by the same arguments in the proof of Theorem 2 of James *et al.* [22], page 96. Thus, the proof is complete.

A.2. Proof of Theorem 3.2

Following the proof of Theorem 3.1 from the beginning to (A.3) with chosen $h'(x) = e^{-\gamma x}$. Using the facts from (A.5) and (A.6), the distribution (A.3) becomes

$$\frac{e^{-\gamma(\mu(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)}}{(\mu(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)^{n+q}} \mathcal{P}_{\tilde{\mu}}(d\mu) \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k) s_k^{n_k} \rho_{Y_k}(ds_k). \quad (\text{A.7})$$

This is the joint distribution of $\tilde{\mu}$, \mathbf{p}_n and Y_k for $k = 1, \dots, \aleph(\mathbf{p}_n)$. Then we apply the gamma identity (3.6) on the term $(\mu(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)^{-(n+q)}$ in (A.7)

$$\frac{1}{(\mu(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)^{n+q}} = \frac{1}{\Gamma(n+q)} \int_{\mathbb{R}^+} e^{-(\mu(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)u} u^{n+q-1} du, \quad (\text{A.8})$$

Here the augmented variable U_n is introduced due to the gamma identity. Incorporating (A.7) with (A.8) and omitting the integrals, it turns out that the following is the joint distribution of U_n , $\tilde{\mu}$, \mathbf{p}_n and Y_k for $k = 1, \dots, \aleph(\mathbf{p}_n)$,

$$e^{-(\gamma+u)\mu(\mathbb{X})} \mathcal{P}_{\tilde{\mu}}(d\mu) \frac{u^{n+q-1}}{\Gamma(n+q)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k) s_k^{n_k} e^{-s_k(\gamma+u)} \rho_{Y_k}(ds_k) du. \quad (\text{A.9})$$

The disintegration between terms in $e^{-(\gamma+u)\mu(\mathbb{X})}$ and $\mathcal{P}_{\tilde{\mu}}(d\mu)$ in (A.9) yields $e^{-\psi_0(\gamma+u)}$ and $\mathcal{P}_{\tilde{\mu}^{(u)}}(d\mu)$ where ψ_0 is defined in (2.9) and the completely random measure $\tilde{\mu}^{(u)}$ has the intensity measure $\nu^{(u)}(dx, ds) = \eta(dx) e^{-(\gamma+u)s} \rho_x(ds)$. Then, (A.9) turns out to be

$$\mathcal{P}_{\tilde{\mu}^{(u)}}(d\mu) e^{-\psi_0(\gamma+u)} \frac{u^{n+q-1}}{\Gamma(n+q)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k) s_k^{n_k} e^{-s_k(\gamma+u)} \rho_{Y_k}(ds_k) du. \quad (\text{A.10})$$

So, conditional on U_n and \mathbf{X}_n , the process $\tilde{\mu}^{(U_n)} + \sum_{k=1}^{\aleph(\mathbf{p}_n)} J_k^{(U_n, \mathbf{X}_n)} \delta_{Y_k}$ is a completely random measure (2.1) (see also Daley and Vere-Jones [8], page 79, Theorem 10.1.III) such that

$$\tilde{\mu}^{(U_n, \mathbf{X}_n)}(dx) \stackrel{d}{=} \tilde{\mu}^{(U_n)}(dx) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} J_k^{(U_n, \mathbf{X}_n)} \delta_{Y_k}(dx), \quad (\text{A.11})$$

where each $J_k^{(U_n, \mathbf{X}_n)}$ has the density $\mathbb{P}(J_k^{(U_n, \mathbf{X}_n)} \in ds) = \frac{s^{n_k} e^{-s(\gamma+U_n)} \rho_{Y_k}(ds)}{\int_{\mathbb{R}^+} s^{n_k} e^{-s(\gamma+U_n)} \rho_{Y_k}(ds)}$ for $k = 1, \dots, \aleph(\mathbf{p}_n)$ and $\{J_1^{(U_n, \mathbf{X}_n)}, \dots, J_{\aleph(\mathbf{p}_n)}^{(U_n, \mathbf{X}_n)}\}$ are conditionally independent. Furthermore, conditional on U_n , $\tilde{\mu}^{(U_n)}$ and $\{J_1^{(U_n, \mathbf{X}_n)}, \dots, J_{\aleph(\mathbf{p}_n)}^{(U_n, \mathbf{X}_n)}\}$ are independent. Lastly the distributional identity $\tilde{G}^{(U_n, \mathbf{X}_n)}(\cdot) \stackrel{d}{=} \tilde{\mu}^{(U_n, \mathbf{X}_n)}(\cdot) / \tilde{\mu}^{(U_n, \mathbf{X}_n)}(\mathbb{X})$ is obtained by the same arguments in the proof of Theorem 2 of James *et al.* [22], page 96, and $\tilde{G}^{(U_n, \mathbf{X}_n)}$ is a normalized completely random measure. Thus, the proof is complete.

A.3. Proof of Proposition 3.1

Following the definition of the predictive distribution,

$$\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\} = \int_{\mathbb{R}^+} E[\tilde{G}^{(U_n, \mathbf{X}_n)}(dx) | \mathbf{X}_n, U_n = u] f_{U_n | \mathbf{X}_n}(u) du, \tag{A.12}$$

and using the result from Theorem 3.2, the expectation (A.12) inside the integral is given by

$$\begin{aligned} & E[\tilde{G}^{(U_n, \mathbf{X}_n)}(dx) | \mathbf{X}_n, U_n] \\ &= E\left[\frac{\tilde{\mu}^{(U_n)}(dx)}{\tilde{\mu}^{(U_n)}(\mathbb{X}) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} J_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n\right] \\ &+ \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} E\left[\frac{J_k^{(U_n, \mathbf{X}_n)}}{\tilde{\mu}^{(U_n)}(\mathbb{X}) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} J_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n\right] \delta_{Y_k}(dx). \end{aligned} \tag{A.13}$$

We use Theorem 3.2 to obtain the explicit results of the expectations (A.13) according to the conditional distributions of $\tilde{\mu}^{(U_n)}$ and $J_k^{(U_n, \mathbf{X}_n)}$ for $k = 1, \dots, \mathfrak{N}(\mathbf{p}_n)$. There are two expected values (A.13) considered in the following. Firstly, an exponential identity in the first integral of (A.13) yields,

$$\begin{aligned} & E\left[\frac{\tilde{\mu}^{(U_n)}(dx)}{\mu^{(U_n)}(\mathbb{X}) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} J_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n\right] \\ &= \int_{\mathbb{R}^+} E\left[\tilde{\mu}^{(U_n)}(dx) e^{-z\tilde{\mu}^{(U_n)}(\mathbb{X})} \prod_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} e^{-zJ_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n\right] dz. \end{aligned} \tag{A.14}$$

We apply a change of measure on $\tilde{\mu}^{(U_n)}$ and compute directly on the expectation with respect to $J_k^{(U_n, \mathbf{X}_n)}$ for $k = 1, \dots, \mathfrak{N}(\mathbf{p}_n)$, (A.14) turns out to be

$$\eta(dx) \times \int_{\mathbb{R}^+} e^{-\psi_{\gamma+U_n}(z)} \tau_{1,x}(\gamma + z + U_n) \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{\tau_{n_k, Y_k}(\gamma + z + U_n)}{\tau_{n_k, Y_k}(\gamma + U_n)} dz. \tag{A.15}$$

Marginalizing over U_n on (A.15) is given by

$$\begin{aligned} & \int_{\mathbb{R}^+} E\left[\frac{\tilde{\mu}^{(U_n)}(dx)}{\tilde{\mu}^{(U_n)}(\mathbb{X}) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} J_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n\right] f_{U_n | \mathbf{X}_n}(u) du \\ &= \eta(dx) \times \int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{-\psi_{\gamma+u}(z)} \tau_{1,x}(\gamma + z + u) \\ &\quad \times \frac{e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_k, Y_k}(\gamma + z + u) u^{n+q-1}}{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_k, Y_k}(\gamma + u) u^{n+q-1} du} du dz. \end{aligned} \tag{A.16}$$

Taking the equality $\psi_0(\gamma + z + u) = \psi_{\gamma+u}(z) + \psi_0(\gamma + u)$ and transforming the upper integral with $(w, v) = (u + z, u)$, (A.16) becomes

$$\begin{aligned} \eta(dx) &\times \frac{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+w)} \tau_{1,x}(\gamma + w) \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_k, Y_k}(\gamma + w) w^{n+q} / (n+q) dw}{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_k, Y_k}(\gamma + u) u^{n+q-1} du} \\ &= \eta(dx) \times \frac{1}{n+q} \int_{\mathbb{R}^+} \tau_{1,x}(\gamma + u) u f_{U_n | \mathbf{X}_n}(u) du. \end{aligned}$$

Now consider the second expectation of (A.13). As before, an exponential identity in the second integral yields,

$$\begin{aligned} E \left[\frac{J_k^{(U_n, \mathbf{X}_n)}}{\mu^{(U_n)}(\mathbb{X}) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} J_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n \right] \\ = \int_{\mathbb{R}^+} E \left[J_k^{(U_n, \mathbf{X}_n)} e^{-z\mu^{(U_n)}(\mathbb{X})} \prod_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} e^{-zJ_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n \right] dz. \end{aligned} \tag{A.17}$$

By direct computation, (A.17) turns out to be

$$\int_{\mathbb{R}^+} e^{-\psi_{\gamma+U_n}(z)} \frac{\tau_{n_k+1, Y_k}(\gamma + z + U_n)}{\tau_{n_k, Y_k}(\gamma + U_n)} \prod_{k'=1, k' \neq k}^{\mathfrak{N}(\mathbf{p}_n)} \frac{\tau_{n_{k'}, Y_{k'}}(\gamma + z + U_n)}{\tau_{n_{k'}, Y_{k'}}(\gamma + U_n)} dz. \tag{A.18}$$

Marginalizing over U_n on (A.18) is given by

$$\begin{aligned} \int_{\mathbb{R}^+} E \left[\frac{J_k^{(U_n, \mathbf{X}_n)}}{\tilde{\mu}^{(U_n)}(\mathbb{X}) + \sum_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} J_{k'}^{(U_n, \mathbf{X}_n)}} \middle| \mathbf{X}_n, U_n \right] f_{U_n | \mathbf{X}_n}(u) du \\ = \int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{-\psi_{\gamma+u}(z)} \tau_{n_k+1, Y_k}(\gamma + z + u) \\ \times \frac{e^{-\psi_0(\gamma+u)} \prod_{k'=1, k' \neq k}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_{k'}, Y_{k'}}(\gamma + z + u) u^{n+q-1}}{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_{k'}, Y_{k'}}(\gamma + u) u^{n+q-1} du} du dz. \end{aligned} \tag{A.19}$$

Taking the equality $\psi_0(\gamma + z + u) = \psi_{\gamma+u}(z) + \psi_0(\gamma + u)$ and transforming the upper integral with $(w, v) = (u + z, u)$, (A.19) becomes

$$\begin{aligned} \frac{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+w)} \tau_{n_k+1, Y_k}(\gamma + w) \prod_{k'=1, k' \neq k}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_{k'}, Y_{k'}}(\gamma + w) w^{n+q} / (n+q) dw}{\int_{\mathbb{R}^+} e^{-\psi_0(\gamma+u)} \prod_{k'=1}^{\mathfrak{N}(\mathbf{p}_n)} \tau_{n_{k'}, Y_{k'}}(\gamma + u) u^{n+q-1} du} \\ = \frac{1}{n+q} \int_{\mathbb{R}^+} \frac{\tau_{n_k+1, Y_k}(\gamma + u)}{\tau_{n_k, Y_k}(\gamma + u)} u f_{U_n | \mathbf{X}_n}(u) du. \end{aligned} \tag{A.20}$$

We conclude that the predictive distribution for X_{n+1} given $\mathbf{X}_n = \{X_1, \dots, X_n\}$ is given by

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\} &= \frac{1}{n+q} \int_{\mathbb{R}^+} u \tau_{1,x}(\gamma+u) f_{U_n | \mathbf{X}_n}(u) du \times \eta(dx) \\ &+ \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{1}{n+q} \int_{\mathbb{R}^+} u \frac{\tau_{n_k+1, Y_k}(\gamma+u)}{\tau_{n_k, Y_k}(\gamma+u)} f_{U_n | \mathbf{X}_n}(u) du \times \delta_{Y_k}(dx). \end{aligned} \tag{A.21}$$

This representation is analogous to James *et al.* [22]. In fact, we prefer an intuitive representation for this predictive distribution for the next proof of the conditional case. A representation of the predictive distribution is given by

$$\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\} = \frac{\omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(x)}{\phi(\mathbf{X}_n)} \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{\omega_{n,k}(Y_k)}{\phi(\mathbf{X}_n)} \delta_{Y_k}(dx), \tag{A.22}$$

where

$$\begin{aligned} \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(x) &= \int_{\mathbb{R}^+} u \tau_{1,x}(\gamma+u) f_{U_n | \mathbf{X}_n}(u) du, \\ \omega_{n,k}(Y_k) &= \int_{\mathbb{R}^+} u \frac{\tau_{n_k+1, Y_k}(\gamma+u)}{\tau_{n_k, Y_k}(\gamma+u)} f_{U_n | \mathbf{X}_n}(u) du, \\ \phi(\mathbf{X}_n) &= \int_{\mathbb{X}} \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(x) \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \omega_{n,k}(Y_k) = n+q. \end{aligned}$$

Clearly, we could write (A.21) or (A.22) as an expectation with respect to the distribution U_n given \mathbf{X}_n , $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\} = E[\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, U_n\} | \mathbf{X}_n]$. However, the term $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, U_n\}$ might not be a proper distribution. We suggest the following expectation instead with respect to the distribution \tilde{U}_n given \mathbf{X}_n , $\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n\} = E[\mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} | \mathbf{X}_n]$. So conditional on \tilde{U}_n , the predictive distribution is given by

$$\begin{aligned} \mathbb{P}\{X_{n+1} \in dx | \mathbf{X}_n, \tilde{U}_n\} &= \frac{\omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(\tilde{U}_n, x)}{\phi(\tilde{U}_n, \mathbf{X}_n)} \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \frac{\omega_{n,k}(\tilde{U}_n, Y_k)}{\phi(\tilde{U}_n, \mathbf{X}_n)} \delta_{Y_k}(dx), \\ \phi(u, \mathbf{X}_n) &= \int_{\mathbb{X}} \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(u, x) \eta(dx) + \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} \omega_{n,k}(u, Y_k), \\ \omega_{n, \mathfrak{N}(\mathbf{p}_n)+1}(u, x) &= u \tau_{1,x}(\gamma+u), \\ \omega_{n,k}(u, Y_k) &= u \frac{\tau_{n_k+1, Y_k}(\gamma+u)}{\tau_{n_k, Y_k}(\gamma+u)}, \quad k = 1, \dots, \mathfrak{N}(\mathbf{p}_n) \quad \text{and} \\ f_{\tilde{U}_n | \mathbf{X}_n}(u) &= \frac{\phi(u, \mathbf{X}_n)}{E[\phi(U_n, \mathbf{X}_n) | \mathbf{X}_n]} f_{U_n | \mathbf{X}_n}(u). \end{aligned}$$

Lastly, the following equality can be achieved by inspection of (A.21) and (A.22), $\phi(\mathbf{X}_n) = E[\phi(U_n|\mathbf{X}_n)|\mathbf{X}_n] = n + q$ and the following equalities can be obtained according to the definitions, $\omega_{n, \aleph(\mathbf{p}_n)+1}(x) = E[\omega_{n, \aleph(\mathbf{p}_n)+1}(U_n, x)|\mathbf{X}_n]$ and $\omega_{n, k}(Y_k) = E[\omega_{n, k}(U_n, Y_k)|\mathbf{X}_n]$ for $k = 1, \dots, \aleph(\mathbf{p}_n)$.

A.4. Proof of Proposition 3.2

The joint distribution of $\{\mathbf{Y}_{\aleph(\mathbf{p}_n)}, \mathbf{p}_n, U_n\}$ could be simply derived from (A.10) of Theorem 3.2, which is proportional to

$$e^{-\psi_0(\gamma+u)} \frac{u^{n+q-1}}{\Gamma(n+q)} \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k) \tau_{n_k, Y_k}(\gamma+u) du,$$

where $\tau_{m,z}(a)$ is defined in (3.17). Distributional results around these random variables $\{\mathbf{Y}_{\aleph(\mathbf{p}_n)}, \mathbf{p}_n, U_n\}$ are achieved immediately with some simple algebra.

A.5. Proof of equation (6.1)

Starting from the marginal distribution of \mathbf{X}_n which is given by

$$\int_{\mathcal{M}_{\mathbb{X}}} \prod_{i=1}^n \frac{\mu(dx_i)}{\mu(\mathbb{X})} \mathcal{P}_{\tilde{\mu}_t}(d\mu) = \int_{\mathcal{M}_{\mathbb{X}}} \prod_{i=1}^n \mu(dx_i) \frac{h(\mu(\mathbb{X}))\mu(\mathbb{X})^{-n}}{E[h(\mu(\mathbb{X}))]} \mathcal{P}_{\tilde{\mu}}(d\mu). \quad (\text{A.23})$$

One could apply the Fubini theorem following from an application of Lemma 2.2 of James [18], page 8, on (A.23) to yield the joint distribution of $\{\mathbf{Y}_{\aleph(\mathbf{p}_n)}, \mathbf{p}_n\}$,

$$\frac{1}{E[h(\mu(\mathbb{X}))]} \int_{\mathcal{M}_{\mathbb{X}} \times (\mathbb{R}^+)^{\aleph(\mathbf{p}_n)}} \frac{h(\mu(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)}{(\mu(\mathbb{X}) + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)^n} \mathcal{P}_{\tilde{\mu}}(d\mu) \prod_{k=1}^{\aleph(\mathbf{p}_n)} \eta(dY_k) s_k^{n_k} \rho(ds_k). \quad (\text{A.24})$$

Then, in (A.24), marginalizing $\mathbf{Y}_{\aleph(\mathbf{p}_n)}$ and replacing the distribution of the total $\tilde{T} = \tilde{\mu}(\mathbb{X})$ by $f_{\tilde{T}}(t)$, the distribution of \mathbf{p}_n is given by the following without the proportional constant,

$$\int_{(\mathbb{R}^+)^{\aleph(\mathbf{p}_n)} \times \mathbb{R}^+} \frac{h(t + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)}{(t + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)^n} \frac{1}{E[h(\tilde{T})]} f_{\tilde{T}}(t) \prod_{k=1}^{\aleph(\mathbf{p}_n)} s_k^{n_k} \rho(ds_k) dt. \quad (\text{A.25})$$

Considering the transformation $t^* = t + \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k$ and (A.25) becomes

$$\int_{(\mathbb{R}^+)^{\aleph(\mathbf{p}_n)} \times \mathbb{R}^+} \mathbb{I}_{\{t^* - \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k > 0\}} \frac{h(t^*)}{(t^*)^n} \frac{f_{\tilde{T}}(t^* - \sum_{k=1}^{\aleph(\mathbf{p}_n)} s_k)}{E[h(\tilde{T})]} \prod_{k=1}^{\aleph(\mathbf{p}_n)} s_k^{n_k} \rho(ds_k) dt^*. \quad (\text{A.26})$$

Rearranging terms in (A.26) yields

$$\int_{(\mathbb{R}^+)^{\mathfrak{N}(\mathbf{p}_n)} \times \mathbb{R}^+} \mathbb{I}_{\{t^* - \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} s_k > 0\}} \frac{f_{\tilde{T}}(t^* - \sum_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} s_k)}{(t^*)^n f_{\tilde{T}}(t^*)} \prod_{k=1}^{\mathfrak{N}(\mathbf{p}_n)} s_k^{n_k} \rho(ds_k) \frac{h(t^*) f_{\tilde{T}}(t^*)}{E[h(\tilde{T})]} dt^*.$$

Thus, the proof is complete.

Acknowledgements

The author thanks Edward Cripps for the tremendous help in revising this article and Robin Milne for proofreading this manuscript and giving constructive comments on the presentation. Their invaluable comments and supports to the realization of this work are deeply appreciated. The author also thanks the associate editor and two referees for their helpful suggestions concerning the presentation of this article and recommending various improvements in exposition. This research was partially supported by Australian Actuarial Research Grant #9201101267.

References

- [1] Aldous, D.J. (1985). Exchangeability and related topics. In *École d'Été de Probabilités de Saint-Flour, XIII—1983. Lecture Notes in Math.* **1117** 1–198. Berlin: Springer. [MR0883646](#)
- [2] Arratia, R., Barbour, A.D. and Tavaré, S. (2003). *Logarithmic Combinatorial Structures: A Probabilistic Approach. EMS Monographs in Mathematics.* Zürich: European Mathematical Society (EMS). [MR2032426](#)
- [3] Blackwell, D. and MacQueen, J.B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. [MR0362614](#)
- [4] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. in Appl. Probab.* **31** 929–953. [MR1747450](#)
- [5] Cifarelli, D.M. and Regazzini, E. (1979). Considerazioni generali sull'impostazione bayesiana di problemi non parametrici. Le medie associative nel contesto del processo aleatorio di Dirichlet. Parte I. *Riv. Mat. Sci. Econom. Social.* **2** 39–52.
- [6] Cifarelli, D.M. and Regazzini, E. (1979). Considerazioni generali sull'impostazione bayesiana di problemi non parametrici. Le medie associative nel contesto del processo aleatorio di Dirichlet. Parte II. *Riv. Mat. Sci. Econom. Social.* **2** 95–111.
- [7] Cifarelli, D.M. and Regazzini, E. (1990). Distribution functions of means of a Dirichlet process. *Ann. Statist.* **18** 429–442. [Correction *Ann. Statist.* **22** (1994) 1633–1634]. [MR1041402](#)
- [8] Daley, D.J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes. Vol. II: General Theory and Structure*, 2nd ed. *Probability and Its Applications (New York)*. New York: Springer. [MR2371524](#)
- [9] Devroye, L. (1986). *Nonuniform Random Variate Generation*. New York: Springer. [MR0836973](#)
- [10] Doksum, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.* **2** 183–201. [MR0373081](#)
- [11] Dykstra, R.L. and Laud, P. (1981). A Bayesian nonparametric approach to reliability. *Ann. Statist.* **9** 356–367. [MR0606619](#)
- [12] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. [MR0350949](#)

- [13] Griffin, J.E. and Steel, M.F.J. (2006). Order-based dependent Dirichlet processes. *J. Amer. Statist. Assoc.* **101** 179–194. [MR2268037](#)
- [14] Griffiths, R.C. and Lessard, S. (2005). Ewens' sampling formula and related formulae: Combinatorial proofs, extensions to variable population size and applications to ages of alleles. *Theor. Popul. Biol.* **68** 167–177.
- [15] Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73** 387–396. [MR0855898](#)
- [16] Ishwaran, H. and James, L.F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* **96** 161–173. [MR1952729](#)
- [17] Ishwaran, H. and James, L.F. (2003). Generalized weighted Chinese restaurant processes for species sampling mixture models. *Statist. Sinica* **13** 1211–1235. [MR2026070](#)
- [18] James, L.F. (2002). Poisson process partition calculus with applications to exchangeable models and Bayesian nonparametrics. Preprint. Available at <http://arxiv.org/abs/math/0205093>.
- [19] James, L.F. (2005). Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages. *Ann. Statist.* **33** 1771–1799. [MR2166562](#)
- [20] James, L.F. (2005). Functionals of Dirichlet processes, the Cifarelli–Regazzini identity and beta-gamma processes. *Ann. Statist.* **33** 647–660. [MR2163155](#)
- [21] 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](#)
- [22] James, L.F., Lijoi, A. and Prünster, I. (2009). Posterior analysis for normalized random measures with independent increments. *Scand. J. Stat.* **36** 76–97. [MR2508332](#)
- [23] Kallenberg, O. (1983). *Random Measures*, 3rd ed. Berlin: Akademie-Verlag. [MR0818219](#)
- [24] Kallenberg, O. (2001). *Foundations of Modern Probability*, 2nd ed. New York: Springer.
- [25] Kingman, J.F.C. (1967). Completely random measures. *Pacific J. Math.* **21** 59–78. [MR0210185](#)
- [26] Kingman, J.F.C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. New York: Oxford Univ. Press. [MR1207584](#)
- [27] Kingman, J.F.C., Taylor, S.J., Hawkes, A.G., Walker, A.M., Cox, D.R., Smith, A.F.M., Hill, B.M., Burville, P.J. and Leonard, T. (1975). Random discrete distribution. *J. Roy. Statist. Soc. Ser. B* **37** 1–22. With a discussion by S. J. Taylor, A. G. Hawkes, A. M. Walker, D. R. Cox, A. F. M. Smith, B. M. Hill, P. J. Burville, T. Leonard and a reply by the author. [MR0368264](#)
- [28] Lau, J.W. and Cripps, E. (2012). Bayesian non-parametric mixtures of GARCH(1, 1) models. *J. Probab. Stat.* **2012** Art. ID 167431. [MR2949443](#)
- [29] Lau, J.W. and So, M.K.P. (2008). Bayesian mixture of autoregressive models. *Comput. Statist. Data Anal.* **53** 38–60. [MR2528591](#)
- [30] Lau, J.W. and So, M.K.P. (2011). A Monte Carlo Markov chain algorithm for a class of mixture time series models. *Stat. Comput.* **21** 69–81. [MR2746604](#)
- [31] Lijoi, A., Mena, R.H. and Prünster, I. (2005). Bayesian nonparametric analysis for a generalized Dirichlet process prior. *Stat. Inference Stoch. Process.* **8** 283–309. [MR2177315](#)
- [32] Lijoi, A., Mena, R.H. and Prünster, I. (2005). Hierarchical mixture modeling with normalized inverse-Gaussian priors. *J. Amer. Statist. Assoc.* **100** 1278–1291. [MR2236441](#)
- [33] Lijoi, A., Mena, R.H. and Prünster, I. (2007). Bayesian nonparametric estimation of the probability of discovering new species. *Biometrika* **94** 769–786. [MR2416792](#)
- [34] Lijoi, A., Mena, R.H. and Prünster, I. (2007). Controlling the reinforcement in Bayesian non-parametric mixture models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **69** 715–740. [MR2370077](#)
- [35] Lijoi, A. and Prünster, I. (2010). Models beyond the Dirichlet process. In *Bayesian Nonparametrics* (N.L. Hjort, C. Holmes, P. Müller and S.G. Walker, eds.). *Camb. Ser. Stat. Probab. Math.* 80–136. Cambridge: Cambridge Univ. Press. [MR2730661](#)

- [36] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12** 351–357. [MR0733519](#)
- [37] Lo, A.Y. (1991). A characterization of the Dirichlet process. *Statist. Probab. Lett.* **12** 185–187. [MR1130354](#)
- [38] Lo, A.Y., Brunner, L.J. and Chan, A.T. (1996). Weighted Chinese restaurant processes and Bayesian mixture models. Research report, Hong Kong Univ. Science and Technology. Available at <http://www.utstat.utoronto.ca/~brunner/papers/wcr96.pdf>.
- [39] 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](#)
- [40] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102** 145–158. [MR1337249](#)
- [41] Pitman, J. (2003). Poisson–Kingman partitions. In *Statistics and Science: A Festschrift for Terry Speed. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **40** 1–34. Beachwood, OH: IMS. [MR2004330](#)
- [42] Pitman, J. (2006). *Combinatorial Stochastic Processes. Lecture Notes in Math.* **1875**. Berlin: Springer. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. [MR2245368](#)
- [43] Pitman, J. and Yor, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900. [MR1434129](#)
- [44] Regazzini, E., Lijoi, A. and Prünster, I. (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31** 560–585. Dedicated to the memory of Herbert E. Robbins. [MR1983542](#)
- [45] Wolpert, R.L. and Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika* **85** 251–267. [MR1649114](#)

Received October 2010 and revised May 2012