*Journal of Nonparametric Statistics*
Vol. 20, No. 5, July 2008, 447–457





# Posterior analysis for some classes of nonparametric models

Antonio Lijoi[a,b], Igor Prünster[c] and S.G. Walker[d]*

[a]*Dipartimento di Economia Politica e Metodi Quantitativi, Università degli Studi di Pavia, Pavia, Italy;* [b]*CNR-IMATI, Milano, Italy;* [c]*Dipartimento di Statistica e Matematica Applicata, Collegio Carlo Alberto and ICER, Università degli Studi di Torino, Torino, Italy;* [d]*Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Kent, UK*

(*Received 2006; final version received 08 May 2008*)

Recently, James [L.F. James, *Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages*, Ann. Statist. 33 (2005), pp. 1771–1799.] and [L.F. James, *Poisson calculus for spatial neutral to the right processes*, Ann. Statist. 34 (2006), pp. 416–440.] has derived important results for various models in Bayesian nonparametric inference. In particular, in ref. [L.F. James, *Poisson calculus for spatial neutral to the right processes*, Ann. Statist. 34 (2006), pp. 416–440.] a spatial version of neutral to the right processes is defined and their posterior distribution derived. Moreover, in ref. [L.F. James, *Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages*, Ann. Statist. 33 (2005), pp. 1771–1799.] the posterior distribution for an intensity or hazard rate modelled as a mixture under a general multiplicative intensity model is obtained. His proofs rely on the so-called Bayesian Poisson partition calculus. Here we provide alternative proofs based on a different technique.

**Keywords:** Bayesian nonparametrics; completely random measure; hazard rate; multiplicative intensity model; neutral to the right prior

*AMS Subject Classification*: 60G57; 62F15; 60G55; 62G05

## 1. Introduction

Recently, James [1] introduced a technique, called Bayesian Poisson partition calculus, which allows the derivation of posterior distributions for a large variety of Bayesian nonparametric models. His technique, whose roots lie in refs. [2,3], consists of a Laplace functional change of measure combined with a Poisson Palm/Fubini calculus on random partitions of the positive integers. See ref. [4] and references therein for an exhaustive account on exchangeable random partitions and applications in areas not directly related to Bayesian nonparametrics.

In ref. [5], the multiplicative intensity model of Aalen [6] is considered. It is well known that the multiplicative intensity model covers a large variety of important applied models such as the simple life testing model, the Cox proportional hazards regression model, the multiple decrement model and Poisson process spatial regression models, among others. A typical Bayesian nonparametric approach, in this area, relies on designing the intensity or hazard function as a mixture with respect

*Corresponding author. Email: S.G.Walker@kent.ac.uk

ISSN 1048-5252 print/ISSN 1029-0311 online

DOI: 10.1080/10485250802196364
http://www.informaworld.com

to gamma (or allied) random measures [3,7–11]. James [5] derived the posterior distribution for general multiplicative intensity models in which the intensity or hazard rate is a mixture driven by any completely random measure (CRM), thus generalising all previous posterior representations. Based upon this result and suitable simulation algorithms, practitioners now have the possibility not only of selecting an appropriate kernel but also of deciding which random measure to adopt.

In James [12], a spatial version of the popular neutral to the right (NTR) processes, termed spatial neutral to the right (SPNTR) processes is introduced. Indeed, NTR priors, due to Doksum [13], have been successfully exploited in the context of survival analysis leading to Bayesian nonparametric analogs of the Kaplan–Meier estimator. See, among others, refs. [14–17]. One of the main drawbacks of NTR priors is represented by the fact that they can be defined only on the real line and not on multi-dimensional or general abstract spaces. The notion of SPNTR prior obviates this. Moreover, it relates in a nice way the literature on NTR processes with the other line of research, initiated by Hjort [18], which consists in modelling the cumulative hazard by means of a suitable CRM, namely the beta process. Within this framework, we recall, *e.g.* [19–21]. James [5], applying Poisson partition calculus, derived the posterior distribution for a general SPNTR prior, again opening up the possibility of exploiting concretely many different alternatives in several important applications.

In this paper, we provide alternative proofs of two results of James [5,12] regarding the posterior distribution of SPNTR priors and mixture priors for multiplicative intensity models by means of a different technique. Given the importance of the posterior characterisations in refs. [5,12], it is useful to have different derivations of them, and this could then set a basis for obtaining posterior distributions also in other models involving CRM. Our approach consists in reading a suitably transformed version of the data 'likelihood' as a derivative of the Laplace functional of the CRM upon which the model is built, setting up a recursion and obtaining the posterior Laplace functional in the limit. Such a device has been, at least to authors' knowledge, first employed in refs. [22,23]. Recently, the posterior distribution of normalised random measures with independent increments, a class of priors introduced in ref. [24], has been derived in ref. [25]. In that paper, in addition to a proof based on the Poisson process calculus technique, one relying on the approach of the present paper is provided as well. Finally, it is worth mentioning that the techniques we adopt are connected to some results obtained in ref. [26], where the authors exploit Faa di Bruno's formula and deduce generalisations of Dobinski's formula.

In Section 2, we recall the definition of CRM and some useful notation. In Sections 3 and 4, we provide alternative proofs of the posterior characterisations of SPNTR models and of multiplicative hazard models, respectively.

## 2. Preliminaries and notation

At the heart of most nonparametric models, there is the concept of CRM introduced by Kingman [27], which we briefly recall here. It is worth noting that an increasing additive process (or increasing Lévy process with not necessarily stationary increments) can always be seen as the càdlàg distribution function induced by a CRM on $\mathbb{R}$.

Consider a measure space $(\mathbb{X}, \mathcal{X})$, where $\mathbb{X}$ is a complete and separable metric space, and $\mathcal{X}$ is the usual Borel $\sigma$-field. For notational simplicity, set $\mathbb{S} = \mathbb{R}^+ \times \mathbb{X}$, and let $\mathcal{S}$ denote the product $\sigma$-algebra $\mathcal{B}(\mathbb{R}^+) \otimes \mathcal{X}$ on $\mathbb{S}$, where, as usual, $\mathcal{B}(\mathbb{R}^+)$ stands for the class of Borel subsets of $\mathbb{R}^+$. Introduce, now, a *Poisson random measure* $N$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in the set of non-negative counting measures on $(\mathbb{S}, \mathcal{S})$, with *intensity measure* $\nu$, *i.e.* $\mathbb{E}[N(\mathrm{d}v, \mathrm{d}x)] = \nu(\mathrm{d}v, \mathrm{d}x)$. Hence, for any $A \in \mathcal{S}$ such that $\nu(A) < \infty$, $N(A)$ is a Poisson random variable of parameter $\nu(A)$ and, given any finite collection of pairwise disjoint sets,

$A_1, \ldots, A_k$, in $\mathcal{S}$, the random variables $N(A_1), \ldots, N(A_k)$ are mutually independent. Throughout the paper, $\mathbb{E}[\cdot]$ will denote expectation with respect to $\mathbb{P}$. Moreover, the intensity measure $\nu$ must satisfy $\int_{\mathbb{R}^+}(v \wedge 1)\nu(\mathrm{d}v, \mathbb{X}) < \infty$, where $a \wedge b = \min\{a, b\}$. See ref. [28] for an exhaustive account on Poisson random measures.

Now, let $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$ be the space of boundedly finite measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. We suppose that $\mathbb{M}$ is equipped with the topology of vague convergence, and $\mathcal{B}(\mathbb{M})$ is the corresponding Borel $\sigma$-algebra. Let $\mu$ be a random element defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$. It is further assumed that $\mu$ can be represented as a linear functional of the Poisson random measure $N$ as follows:

$$\mu(B) = \int_{\mathbb{R}^+ \times B} g(v) N(\mathrm{d}v, \mathrm{d}x) \quad \text{for any } B \in \mathcal{X},$$

where $g : \mathbb{X} \to \mathbb{R}^+$ is some measurable function. It can be easily seen from the properties of $N$ that $\mu$ is, in the terminology of ref. [27], a CRM on $\mathbb{X}$, *i.e.* for any collection of disjoint sets $B_1, B_2, \ldots$ in $\mathcal{X}$, the random variables $\mu(B_1), \mu(B_2), \ldots$ are mutually independent.

Let, now, $\mathcal{H}_{\nu,g}$ be the space of functions $h : \mathbb{X} \to \mathbb{R}^+$ such that $\int_{\mathbb{S}}[1 - \mathrm{e}^{-g(v)\,h(x)}]\nu(\mathrm{d}v, \mathrm{d}x)$ is finite. Then, $\mu$ is uniquely characterised by its *Laplace functional* which, for any $h$ in $\mathcal{H}_{\nu,g}$, is given by

$$\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{X}} h(x)\,\mu(\mathrm{d}x)}\right] = \mathrm{e}^{-\int_{\mathbb{S}}[1-\mathrm{e}^{-g(v)\,h(x)}]\nu(\mathrm{d}v,\mathrm{d}x)} =: \mathrm{e}^{-\psi_{\nu,g}(h)}. \tag{1}$$

When $g(v) \equiv v$, we write $\mathcal{H}_{\nu,g} = \mathcal{H}_\nu$ and $\psi_{\nu,g} = \psi_\nu$. See ref. [29] for details and further references on CRMs.

## 3. Posterior analysis of SPNTR models

Let us start by recalling the precise definition of a SPNTR random probability measure given in ref. [12]. To this end, we set $\mathbb{W} = [0, 1] \times \mathbb{S}$, with $\mathcal{W} = \mathcal{B}([0, 1]) \otimes \mathcal{S}$, and introduce a Poisson random measure $N$ on $(\mathbb{W}, \mathcal{W})$ with intensity measure of the form

$$\nu(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) = \rho(\mathrm{d}v|s)\Lambda_0(\mathrm{d}s, \mathrm{d}z). \tag{2}$$

In Equation (2), $\rho$ is a Lévy density and $\Lambda_0$ is a hazard measure on $(\mathbb{S}, \mathcal{S})$. Based on a Poisson random measure with intensity of the form (2), define the following CRMs:

$$\Lambda(C) = \int_{[0,1]\times C} vN(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) \quad \text{for any } C \in \mathcal{S}, \tag{3}$$

$$Z(C) = \int_{[0,1]\times C} -\log(1 - v)\, N(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) \quad \text{for any } C \in \mathcal{S}. \tag{4}$$

We are now in a position to recall the James' definition of SPNTR random probability measure.

Definition [12] *Given $\Lambda$ and $Z$ defined as in Equations* (3) *and* (4)*, respectively, a SPNTR process is a random probability measure on* $(\mathbb{S}, \mathcal{S})$ *defined by means of the relation*

$$F(\mathrm{d}t, \mathrm{d}x) \overset{\mathrm{d}}{=} \mathrm{e}^{-Z(A_t)}\Lambda(\mathrm{d}t, \mathrm{d}x),$$

*where* $A_t = (0, t) \times \mathbb{X}$.

Note that by integrating over $\mathbb{X}$, one obtains the usual NTR process [13] on $\mathbb{R}^+$, *i.e.* $F(\mathrm{d}t) = \mathrm{e}^{-Z(A_t)}\Lambda(\mathrm{d}t, \mathbb{X})$.

We now move to considering the posterior distribution of a SPNTR model. We first remark that a SPNTR measure is almost surely discrete, which can be seen, *e.g.*, from ref. [30]; this implies that samples from a SPNTR measure will contain ties with positive probability. Consequently, in deriving the posterior, one has to consider the case of samples of size $n$ for which only $k \leq n$ observations are distinct. Let $(\boldsymbol{T}, \boldsymbol{X}) = \{(T_1, X_1), \ldots, (T_n, X_n)\}$ be a sample of size $n$ and denote by $(\boldsymbol{T}^*, \boldsymbol{X}^*) = \{(T_1^*, X_1^*), \ldots, (T_k^*, X_k^*)\}$ the $k \leq n$ different observations with frequencies $n_1, \ldots, n_k$, respectively, and agree that the pairs in $(\boldsymbol{T}^*, \boldsymbol{X}^*)$ are set in an increasing order with respect to the first coordinate, *i.e.* $(T_i^*, X_i^*)$ and $(T_j^*, X_j^*)$ are such that $T_i^* < T_j^*$ for any $i < j$. Moreover, define $\bar{n}_j = \sum_{i=j}^k n_i$ and $\bar{N}(s) = \sum_{i=1}^k \bar{n}_i \mathbb{I}_{[T_{i-1}^*, T_i^*)}(s)$, where we agree on $T_0^* = 0$.

PROPOSITION 1 [12, Proposition 4.1] *Let $F(\mathrm{d}s, \mathrm{d}z) = \mathrm{e}^{-Z(A_s)} \Lambda(\mathrm{d}s, \mathrm{d}z)$ be a SPNTR process. The posterior distribution of $F$, given $(\boldsymbol{T}, \boldsymbol{X})$, is a SPNTR process defined by*

$$F^{(n)}(\mathrm{d}s, \mathrm{d}z) \overset{d}{=} \mathrm{e}^{-Z^{(n)}(A_s)} \Lambda^{(n)}(\mathrm{d}s, \mathrm{d}z) \tag{5}$$

*with the following specifications*:

(i) $\Lambda^{(n)}(\mathrm{d}s, \mathrm{d}z) \overset{d}{=} \Lambda^*(\mathrm{d}s, \mathrm{d}z) + \sum_{i=1}^k J_i \delta_{(T_i^*, X_i^*)}(\mathrm{d}s, \mathrm{d}z)$, *where* $\Lambda^*(\mathrm{d}s, \mathrm{d}z) = \int_{[0,1]} v N^*(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z)$ *is a CRM with Poisson intensity*

$$\nu^*(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z) = (1 - v)^{\bar{N}(s)} \rho(\mathrm{d}v|s) \Lambda_0(\mathrm{d}s, \mathrm{d}z), \tag{6}$$

*and for $i = 1, \ldots, k$, $(T_i^*, X_i^*)$ is a fixed point of discontinuity with corresponding jump $J_i$ distributed as*

$$f_{J_i}(\mathrm{d}v) = \frac{v^{n_i}(1 - v)^{\bar{n}_{i+1}} \rho(\mathrm{d}v|T_i^*)}{\int_{[0,1]} v^{n_i}(1 - v)^{\bar{n}_{i+1}} \rho(\mathrm{d}v|T_i^*)}. \tag{7}$$

*Moreover, the $J_i$'s are conditionally independent of $\Lambda^*$.*

(ii) $Z^{(n)}(\mathrm{d}s, \mathrm{d}z) \overset{d}{=} Z^*(\mathrm{d}s, \mathrm{d}z) + \sum_{i=1}^k K_i \delta_{(T_i^*, X_i^*)}(\mathrm{d}s, \mathrm{d}z)$, *where* $Z^*(\mathrm{d}s, \mathrm{d}z) = \int_{[0,1]} -\log(1 - v) N^*(\mathrm{d}v, \mathrm{d}s, \mathrm{d}z)$ *is a CRM with intensity (6), and for $i = 1, \ldots, k$, $(T_i^*, X_i^*)$ is a fixed point of discontinuity with corresponding jump $K_i \overset{d}{=} -\log(1 - J_i)$, where $J_i$ is distributed as in Equation* (7)*.*

*Finally, Equation* (5) *can be rewritten as*

$$F^{(n)}(\mathrm{d}s, \mathrm{d}z) \overset{d}{=} \mathrm{e}^{-Z^*(A_s)} \prod_{i=1}^k (1 - \mathbb{I}_{(T_i^*, \infty)}(s) J_i) \Lambda^*(\mathrm{d}s, \mathrm{d}z) + \sum_{i=1}^k P_i^* \delta_{(T_i^*, X_i^*)}(\mathrm{d}s, \mathrm{d}z), \tag{8}$$

*where all the quantities are as above, and for $i = 1, \ldots, k$, $P_i^*$ is equal in distribution to $J_i \mathrm{e}^{-Z^*(A_{T_i^*})} \prod_{j=1}^{i-1}(1 - J_j)$.*

*Proof* The proof-strategy relies on the derivation of the posterior Laplace functional of $\Lambda$ defined in Equation (3), which then uniquely characterises its posterior distribution. Given the posterior distribution of $\Lambda$, the other parts of the result follow by simple arguments.

By Equation (1), the Laplace functional of a CRM with intensity (2) is of the form

$$\mathrm{e}^{-\psi_{\nu,g}(h)} := \mathrm{e}^{-\int_{\mathbb{W}} (1 - \mathrm{e}^{-h(s,z)g(v)}) \rho(\mathrm{d}v|s) \Lambda_0(\mathrm{d}s, \mathrm{d}z)}$$

for any $h \in \mathcal{H}_{\nu,g}$. Consider now a set $A^\epsilon(k) \subset \mathcal{S}$ defined as the product set $\times_{i=1}^k A_{i,\epsilon}^{n_i}$ with $A_{i,\epsilon} = (t_i^* - \epsilon, t_i^* + \epsilon) \times B_\epsilon(x_i^*)$, where $B_\epsilon(x_i^*)$ denotes a ball of size $\epsilon$ around the point $x_i^*$. Note that

$(\boldsymbol{T}, \boldsymbol{X}) \in A^{\epsilon}(k)$ corresponds to having $n_i$ observations in $A_{i,\epsilon}$ for $i = 1, \ldots, k$; by letting $\epsilon \downarrow 0$, $(\boldsymbol{T}, \boldsymbol{X}) \in A^{\epsilon}(k)$ reduces to a sample $(\boldsymbol{T}, \boldsymbol{X})$ featuring $k$ distinct values $(t_i^*, x_i^*)$ with frequency $n_i$ for $i = 1, \ldots, k$. Our aim is to derive the posterior Laplace functional of $\Lambda$, *i.e.*

$$\lim_{\epsilon \downarrow 0} \mathbb{E}\left[ \mathrm{e}^{-\int_{\mathbb{S}} h(s,z)\Lambda(\mathrm{d}s,\mathrm{d}z)} \,|(\boldsymbol{T}, \boldsymbol{X}) \in A^{\epsilon}(k) \right]. \tag{9}$$

The conditional expectation in Equation (9), before evaluating the limit, can be expressed as

$$\frac{\mathbb{E}\left[ \mathrm{e}^{-\int_{\mathbb{S}} h(s,z)\Lambda(\mathrm{d}s,\mathrm{d}z)} \prod_{i=1}^{k} \Phi_{i,\epsilon}^{n_i} \right]}{\mathbb{E}\left[ \prod_{i=1}^{k} \Phi_{i,\epsilon}^{n_i} \right]} \tag{10}$$

with $\Phi_{i,\epsilon} := \int_{A_{i,\epsilon}} \mathrm{e}^{Z(A_t)} \Lambda(\mathrm{d}t, \mathrm{d}x)$ denoting the (random) probability that an observation falls in $A_{i,\epsilon}$. Notice that

$$\mathrm{e}^{-Z(A_{t_i^*+\epsilon})} \int_{A_{i,\epsilon}} \Lambda(\mathrm{d}t, \mathrm{d}x) \le \Phi_{i,\epsilon} \le \mathrm{e}^{-Z(A_{t_i^*-\epsilon})} \int_{A_{i,\epsilon}} \Lambda(\mathrm{d}t, \mathrm{d}x). \tag{11}$$

Hence a lower bound for the numerator becomes

$$\mathbb{E}\left[ \mathrm{e}^{-\int_{\mathbb{W}} (h(s,z)\, v - \bar{N}_{\epsilon}(s) \log(1-v))\, N(\mathrm{d}v,\mathrm{d}s,\mathrm{d}z)} \prod_{i=1}^{k} \left( \int_{[0,1]\times A_{i,\epsilon}} v\, N(\mathrm{d}v, \mathrm{d}t, \mathrm{d}x) \right)^{n_i} \right]$$

having set $\bar{N}_{\epsilon}(s) = \sum_{i=1}^{k} \bar{n}_i \mathbb{I}_{[t_{i-1}^*+\epsilon, t_i^*+\epsilon)}(s)$ with $t_0^* = 0$. Now, let

$$g_{\epsilon,\lambda}^{C}(v, s, z) := \mathbb{I}_C(s, z)[h(s, z)\, v - \bar{N}_{\epsilon}(s)\, \log(1 - v) + \lambda\, v],$$

where $\lambda$ is a constant, and $C$ some set in $\mathscr{S}$ and also set $g_{\epsilon,0}^{\mathbb{S}} := g_{\epsilon}$. If $C_{\epsilon} = \cap_{i=1}^{k} A_{i,\epsilon}^{c}$, one can exploit the independence of the increments of $N$ in order to decompose the expected value as

$$\mathrm{e}^{-\psi_v(g_{\epsilon,0}^{C_\epsilon})} \prod_{i=1}^{k} (-1)^{n_i} \frac{\mathrm{d}^{n_i}}{\mathrm{d}\lambda^{n_i}} \mathrm{e}^{-\psi_v(g_{\epsilon,\lambda}^{A_{i,\epsilon}})} \bigg|_{\lambda=0} = \mathrm{e}^{-\psi_v(g_{\epsilon})} \prod_{i=1}^{k} V_{i,\epsilon}^{(n_i)},$$

where

$$V_{i,\epsilon}^{(n_i)} = \mathrm{e}^{\psi_v(g_{\epsilon,0}^{A_{i,\epsilon}})} (-1)^{n_i} \frac{\mathrm{d}^{n_i}}{\mathrm{d}\lambda^{n_i}} \mathrm{e}^{-\psi_v(g_{\epsilon,\lambda}^{A_{i,\epsilon}})} \bigg|_{\lambda=0}.$$

In order to evaluate $V_{i,\epsilon}^{(n_i)}$, one can exploit the following recursive relation:

$$\begin{aligned} V_{i,\epsilon}^{(n_i)} &= \Lambda_0(A_{i,\epsilon}) \sum_{j=0}^{n_i-1} \binom{n_i - 1}{j} \xi_{n_i-j}^{(i,\epsilon)} V_{i,\epsilon}^{(j)} \\ &= \Lambda_0(A_{i,\epsilon})\, \Delta_{i,\epsilon}^{(n_i)} \end{aligned}$$

for any $n_i \ge 1$, where $V_{i,\epsilon}^{(0)} \equiv 1$, and for any $m \ge 1$ and $i \in \{1, \ldots, k\}$,

$$\xi_m^{(i,\epsilon)} = \int_{[0,1]\times A_{i,\epsilon}} v^m\, \mathrm{e}^{-g_{\epsilon}(v,t,x)}\, \rho(\mathrm{d}v|t) \frac{\Lambda_0(\mathrm{d}t, \mathrm{d}x)}{\Lambda_0(A_{i,\epsilon})}.$$

One observes that $\Delta_{i,\epsilon}^{(n_i)} = \xi_{n_i}^{(i,\epsilon)} + K_{i,\epsilon}$, where $K_{i,\epsilon}$ is such that $\lim_{\epsilon\downarrow 0}(K_{i,\epsilon}/\Lambda_0(A_{i,\epsilon})) = K_i < \infty$. Moreover,

$$\lim_{\epsilon\downarrow 0} \xi_{i,\epsilon}^{(n_i)} = \int_{[0,1]} \mathrm{e}^{-h(t_i^*,x_i^*)\,v} v^{n_i}(1-v)^{\bar{n}_{i+1}}\,\rho(\mathrm{d}v|t_i^*) =: \xi_i^{(n_i)}.$$

Hence, the numerator of the lower bound of Equation (10) can be rewritten as follows:

$$\mathrm{e}^{-\psi_v(g_\epsilon)} \prod_{i=1}^{k} \left\{ \Lambda_0(A_{i,\epsilon})\,\xi_{i,\epsilon}^{(n_i)} + o(\Lambda(A_{i,\epsilon})) \right\}, \tag{12}$$

where, as usual, $g(x) = o(f(x))$ as $x \to 0$ means that $\lim_{x\to 0}(g(x)/f(x)) = 0$. On the other hand, by Equation (11), an upper bound for the denominator is

$$\mathbb{E}\left[ \mathrm{e}^{-\int_{\mathbb{W}} -\bar{N}_{-\epsilon}(s)\log(1-v)N(\mathrm{d}v,\mathrm{d}s,\mathrm{d}x)} \prod_{i=1}^{k} \left( \int_{[0,1]\times A_{i,\epsilon}} v\,N(\mathrm{d}v,\mathrm{d}s,\mathrm{d}x) \right)^{n_i} \right], \tag{13}$$

where $\bar{N}_{-\epsilon}(s) := \sum_{i=1}^{k} \bar{n}_i \mathbb{I}_{[t_{i-1}^*-\epsilon,t_i^*-\epsilon)}(s)$. One can then resort to arguments analogous to those leading to the lower bound for the numerator and re-express Equation (13) as

$$\mathrm{e}^{-\psi_v(l_\epsilon)} \prod_{i=1}^{k} \left\{ \Lambda_0(A_{i,\epsilon}) \int_{[0,1]\times A_{i,\epsilon}} v^{n_i}(1-v)^{\bar{N}_{-\epsilon}(s)}\rho(\mathrm{d}v|s)\frac{\Lambda_0(\mathrm{d}s,\mathrm{d}x)}{\Lambda_0(A_{i,\epsilon})} + o(\Lambda(A_{i,\epsilon})) \right\} \tag{14}$$

having set $l_\epsilon(v,s,z) := \bar{N}_\epsilon(s)\,\log(1-v)$. Consider now, the ratio of the lower bound for the numerator given in Equation (12), and the upper bound for the denominator given in Equation (14), and take the limit as $\epsilon \downarrow 0$. This yields a lower bound for the posterior Laplace functional of $\Lambda$ (Equation (9)) coinciding with

$$\mathrm{e}^{-\int_{\mathbb{W}}(1-\mathrm{e}^{-h(s,x)\,v})\,(1-v)^{\bar{N}(s)}\rho(\mathrm{d}v|s)\,\Lambda_0(\mathrm{d}s,\mathrm{d}x)} \prod_{i=1}^{k} \int_{[0,1]} \mathrm{e}^{-h(t_i^*,x_i^*)\,v}\,\frac{v^{n_i}(1-v)^{\bar{n}_{i+1}}\,\rho(\mathrm{d}v|t_i^*)}{\int_{[0,1]} v^{n_i}\,(1-v)^{\bar{n}_{i+1}}\,\rho(\mathrm{d}v|t_i^*)},$$

which agrees with the posterior representation of $\Lambda$ given in Point (i). The same result can be obtained by deriving an upper bound for Equation (9) by means of Equation (11), and letting $\epsilon \downarrow 0$. This completes the proof of Point (i).

As for Point (ii), from the posterior distribution of $\Lambda$, one can easily deduce the posterior distribution of $Z$ and, hence, of the SPNTR measure. Note that the posterior distribution of a linear functional of a random measure $\mu$, conditional on a vector of observations $\boldsymbol{Y} = (Y_1,\ldots,Y_n)$, is given by the linear functional of the posterior distribution of $\mu$. In other terms, $(\int f(y)\mu(\mathrm{d}y)|\boldsymbol{Y}) \stackrel{d}{=} \int f(y)\,\mu^{(n)}(\mathrm{d}y)$, where $\mu^{(n)}$ is a random measure whose distribution coincides with the conditional distribution of $\mu$, given $\boldsymbol{Y}$. Hence, the distribution of the background-driving Poisson measure $N$ in Equation (3), given $(\boldsymbol{T}^*,\boldsymbol{X}^*)$, is of the form $N^* + \sum_{i=1}^{k} \delta_{(J_i,T_i^*,X_i^*)}$, where $N^*$ is a Poisson random measure with intensity (6). Consequently, the posterior distribution of $Z$, defined in terms of $N$ via Equation (4), corresponds to $Z^{(n)}(\mathrm{d}s,\mathrm{d}z) = \int_{[0,1]} -\log(1-v)\,[N^*(\mathrm{d}v,\mathrm{d}s,\mathrm{d}z) + \sum_{i=1}^{k}\delta_{(J_i,T_i^*,X_i^*)}(\mathrm{d}v,\mathrm{d}s,\mathrm{d}z)]$, which leads to the statements in Point (ii). Given $\Lambda^{(n)}$ and $Z^{(n)}$, Equations (5) and (8) now easily follow. ■

## 4. Posterior analysis of multiplicative intensity models

Here, we consider a general model for multiplicative intensities as described, *e.g.* in ref. [5]. Let $\mathbb{X}$ and $\mathbb{Y}$ be Polish spaces endowed with their Borel $\sigma$-algebra $\mathcal{X}$ and $\mathcal{Y}$, respectively. Consider a kernel $K$ on $\mathbb{X} \times \mathbb{Y}$ taking values in $\mathbb{R}^+$ such that $y \mapsto K(C|y)$ is $\mathscr{Y}$-measurable for any $C \in \mathscr{X}$ and, for some $\sigma$-finite measure $\tau$ on $\mathbb{X}$, $\int_{\cdot} K(x, y)\tau(\mathrm{d}x)$ is a $\sigma$-finite measure on $\mathscr{B}(\mathbb{R}^+)$ for any $y$ in $\mathbb{Y}$. Consider, now, the random intensity

$$\lambda(x) = \int_{\mathbb{Y}} K(x, y)\,\mu(\mathrm{d}y),$$

where $\mu$ is a general CRM on $(\mathbb{Y}, \mathscr{Y})$ as defined in Section 2. Write the intensity of $\mu$ as $\nu(\mathrm{d}v, \mathrm{d}y) = \rho(\mathrm{d}v|y)\,\alpha(\mathrm{d}y)$, and assume $\alpha$ is non-atomic. Theorem 4.1 in ref. [5] provides a full description of the posterior distribution of $\mu$ given the observations $\boldsymbol{X} = (X_1, \dots, X_n)$ generated by a multiplicative intensity model. Here, we provide an alternative proof of this result based on techniques analogous to those exploited for the derivation of the posterior distribution of a SPNTR. As in ref. [5], the likelihood function is given by

$$\mathcal{L}(\mu; \boldsymbol{x}) = \mathrm{e}^{-\int_{\mathbb{Y}} g_m(y)\mu(\mathrm{d}y)} \prod_{i=1}^{n} \int_{\mathbb{Y}} K(x_i, y)\,\mu(\mathrm{d}y), \tag{15}$$

where $m \geq n$ denotes the number of observations of which $\boldsymbol{x} = (x_1, \dots, x_n)$ are observable, and the remaining $m - n$ are censored. The function $g_m(y) := \sum_{i=1}^{m} \int_{\mathbb{X}} U_i(x)\,K(x, y)\,\tau(\mathrm{d}x)$ is defined in terms of the predictable and observable processes $\{U_i(x) : x \in \mathbb{X}\}$, $i = 1, \dots, m$. Moreover, the kernel $K(\cdot, \cdot)$ is chosen in such a way that $g_m$ is a non-negative and measurable function with bounded support on $\mathbb{Y}$. See ref. [5] for a discussion on the generality of this model.

If we condition on the latent variables $\boldsymbol{Y} = (Y_1, \dots, Y_n)$, the likelihood above reduces to

$$\mathcal{L}(\mu; \boldsymbol{x}, \boldsymbol{y}) = \mathrm{e}^{-\int_{\mathbb{Y}} g_m(y)\mu(\mathrm{d}y)} \prod_{i=1}^{n} K(x_i; y_i)\mu(\mathrm{d}y_i).$$

The above can be usefully rewritten so as to take into account the fact that the $Y_i$'s may feature some ties. In other terms, $k \leq n$ latent variables are distinct, and we denote their values by $Y_1^*, \dots, Y_k^*$. Correspondingly, one has

$$\mathcal{L}(\mu; \boldsymbol{x}, \boldsymbol{y}) = \mathrm{e}^{-\int_{\mathbb{Y}} g_m(y)\mu(\mathrm{d}y)} \prod_{i=1}^{k} [\mu(\mathrm{d}y_i^*)]^{n_i} \prod_{j \in C_i} K(x_j; y_i^*),$$

where $C_i = \{r : y_r = y_i^*\}$. Before presenting our alternative proof of the posterior characterisation, we introduce some useful notation. Let $\boldsymbol{\Pi}_n = (K_n, \boldsymbol{N}_{K_n})$ be a random vector with the number of classes and the frequencies of each class generated by a random partition of the set of integers $[n] = \{1, \dots, n\}$. In other terms, the realisation $\boldsymbol{\Pi}_n = (k, n_1, \dots, n_k)$ corresponds to a partition of $[n]$ into $k$ sets with respective frequencies $n_1, \dots, n_k$. Clearly, $k \in \{1, \dots, n\}$ and $\boldsymbol{u} = (n_1, \dots, n_k) \in A_{k,n}$, where $A_{k,n} := \{(n_1, \dots, n_k) : n_j \geq 1, \sum_{j=1}^{k} n_j = n\}$. Finally, set

$$\omega_{n_j}(y) = \int_{\mathbb{R}^+} v^{n_j}\,\mathrm{e}^{-v g_m(y)} \rho(\mathrm{d}v|y).$$

Proposition 2 [5, Theorem 4.1] *Let $\lambda(x) = \int_{\mathbb{Y}} K(x, y)\,\mu(\mathrm{d}y)$ be a random intensity on $\mathbb{X}$ with $\nu(\mathrm{d}v, \mathrm{d}y) = \rho(\mathrm{d}v|y)\,\alpha(\mathrm{d}y)$ under the model* (15). *Then, given $\boldsymbol{X}$, the posterior distribution of $\lambda$ can be characterised as follows*:

(i) *Given* $\boldsymbol{Y}$ *and* $\boldsymbol{X}$*, the conditional distribution of* $\mu$ *coincides with the distribution of the random measure*

$$\mu^{(n)} \stackrel{d}{=} \mu^* + \sum_{i=1}^{k} J_i\, \delta_{Y_j^*},$$

*where* $\mu^*$ *is a CRM with intensity* $\nu^*(\mathrm{d}v, \mathrm{d}y) = \mathrm{e}^{-v g_m(y)}\, \rho(\mathrm{d}v|y)\, \alpha(\mathrm{d}y)$*, and for* $i = 1, \ldots, k$*,* $Y_i^*$ *is a fixed point of discontinuity with a corresponding jump* $J_i$ *distributed as*

$$f_{J_i}(\mathrm{d}v) = \frac{v^{n_i}\,\mathrm{e}^{-v g_m(y_i^*)}\rho(\mathrm{d}v|y_i^*)}{\int_{\mathbb{R}^+} v^{n_i}\,\mathrm{e}^{-v g_m(y_i^*)}\rho(\mathrm{d}v|y_i^*)}.$$

*Moreover, the* $J_i$*'s are, conditionally on the* $Y_j^*$*'s, independent from* $\mu^*$*.*

(ii) *Conditionally on* $\boldsymbol{X}$ *and on* $\boldsymbol{\Pi}_n = (k, \boldsymbol{n})$*, the* $Y_j^*$*'s are independent and have a distribution given by*

$$f_{Y_j^*}(\mathrm{d}y) = \frac{\omega_{n_j}(y)\prod_{i\in C_j} K(x_i, y)\alpha(\mathrm{d}y)}{\int_{\mathbb{Y}} \omega_{n_j}(y)\prod_{i\in C_j} K(x_i, y)\alpha(\mathrm{d}y)}, \tag{16}$$

*and, given* $\boldsymbol{X}$*, the conditional probability that* $\boldsymbol{\Pi}_n = (k, \boldsymbol{n})$ *is*

$$\frac{\prod_{j=1}^{k} \int_{\mathbb{Y}} \omega_{n_j}(y)\prod_{i\in C_j} K(x_i, y)\alpha(\mathrm{d}y)}{\sum_{i=1}^{n}\sum_{\boldsymbol{n}\in A_{i,n}}\prod_{j=1}^{i} \int_{\mathbb{Y}} \omega_{n_j}(y)\prod_{i\in C_j} K(x_i, y)\alpha(\mathrm{d}y)}. \tag{17}$$

*Proof* Following arguments analogous to those employed for the evaluation of the posterior distribution for a SPNTR process prior, one can proceed to the determination of the posterior Laplace functional of $\mu$, *i.e.* $\mathbb{E}[\mathrm{e}^{-\mu(f)}|\boldsymbol{X}]$, where $\mu(f) = \int_{\mathbb{Y}} f(y)\,\mu(\mathrm{d}y)$ for any measurable function $f : \mathbb{Y} \to \mathbb{R}^+$ such that $\int_{\mathbb{R}^+\times\mathbb{Y}}(1 - \mathrm{e}^{-v\, f(y)})\rho(\mathrm{d}v|y)\alpha(\mathrm{d}y) < \infty$. To this end, suppose the vector $\boldsymbol{y}$ consists of $k$ distinct observations $y_1^*, \ldots, y_k^*$, and let $y_i = y_{j(i)}^*$, where $j(i) \in \{1, \ldots, k\}$. If $B_\epsilon(x)$ stands for the ball of radius $\epsilon$ around point $x$, denote by $A_i^\epsilon(k)$ the rectangle $B_\epsilon(x_i) \times B_\epsilon(y_{j(i)}^*)$, then

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{-\mu(f)}\,|\boldsymbol{X}, \boldsymbol{Y}\right] &= \lim_{\epsilon\downarrow 0} \mathbb{E}\left[\mathrm{e}^{-\mu(f)}\,\middle|\, (X_i, Y_i) \in A_i^\epsilon(k), i = 1, \ldots, k\right] \\
&= \lim_{\epsilon\downarrow 0} \frac{\mathbb{E}\left[\mathrm{e}^{-\mu(f+g_m)}\prod_{j=1}^{k}\prod_{i\in C_j}\int_{B_\epsilon(x_i)}\int_{B_\epsilon(y_j^*)} K(x, y)\mu(\mathrm{d}y)\tau(\mathrm{d}x)\right]}{\mathbb{E}\left[\mathrm{e}^{-\mu(g_m)}\prod_{j=1}^{k}\prod_{i\in C_j}\int_{B_\epsilon(x_i)}\int_{B_\epsilon(y_j^*)} K(x, y)\mu(\mathrm{d}y)\tau(\mathrm{d}x)\right]}.
\end{aligned}$$

As far as the numerator is concerned, one notes that it can be rewritten as

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{-\int_{\mathbb{Y}^*}(f(y)+g_m(y))\,\mu(\mathrm{d}y)}\right] \\
&\quad\times\prod_{j=1}^{k}\mathbb{E}\left[\mathrm{e}^{-\int_{B_\epsilon(y_j^*)}[f(y)+g_m(y)]\,\mu(\mathrm{d}y)}\prod_{i\in C_j}\int_{B_\epsilon(y_j^*)}\int_{B_\epsilon(x_i)} K(x, y)\tau(\mathrm{d}x)\mu(\mathrm{d}y)\right]
\end{aligned} \tag{18}$$

and $\mathbb{Y}^* = (\cup_{j=1}^{k} B_\epsilon(y_j^*))^c$. For the moment, it is useful to set $\gamma_{i,\epsilon}(y) := \int_{B_\epsilon(x_i)} K(x, y)\tau(\mathrm{d}x)$ for any $y \in \mathbb{Y}$ and $i \in C_j$ and introduce the function $q_{j,\epsilon}^{\boldsymbol{\lambda}} = \left\{f + g_m + \sum_{i=1}^{n_j}\lambda_i\,\gamma_{i,\epsilon}\right\}\mathbb{I}_{B_\epsilon(y_j^*)}$.

Then, recall that $n_j$ is the cardinality of the set of indices and for simplicity, assume that $C_j = \{1, \ldots, n_j\}$. One notes that Equation (18) also coincides with

$$\mathrm{e}^{-\psi_\nu([f+g_m]\mathbb{I}_{\mathbb{Y}^*})} \prod_{j=1}^{k} (-1)^{n_j} \frac{\partial^{n_j}}{\partial\lambda_1 \cdots \partial\lambda_{n_j}} \mathrm{e}^{-\psi_\nu(q^{\boldsymbol\lambda}_{j,\epsilon})}\bigg|_{\boldsymbol\lambda=\mathbf{0}} = \mathrm{e}^{-\psi_\nu(f+g_m)} \prod_{j=1}^{k} V^{(n_j)}_{j,\epsilon},$$

where $\mathbf{0} = (0, \ldots, 0)$ is an $n_j$-dimensional vector, and

$$V^{(n_j)}_{j,\epsilon} = \mathrm{e}^{\psi_\nu(q^{\mathbf{0}}_{j,\epsilon})} (-1)^{n_j} \frac{\partial^{n_j}}{\partial\lambda_1 \cdots \partial\lambda_{n_j}} \mathrm{e}^{-\psi_\nu(q^{\boldsymbol\lambda}_{j,\epsilon})}\bigg|_{\boldsymbol\lambda=\mathbf{0}}.$$

The recursive formula recalled in the proof of Proposition 1 applies in this case as well, yielding

$$V^{(n_j)}_{j,\epsilon} = \alpha(B_\epsilon(y^*_j)) \sum_{i=0}^{n_j-1} \binom{n_j-1}{i} \zeta^{(j,\epsilon)}_{n_j-i} V^{(i)}_{j,\epsilon} = \alpha(B_\epsilon(y^*_j))\, \Delta^{(n_j)}_{j,\epsilon},$$

where $\Delta^{(n_j)}_{j,\epsilon} = \sum_{i=0}^{n_j-1} {}^{n_j-1}_{\;\;i}\, \zeta^{(j,\epsilon)}_{n_j-i} V^{(i)}_{j,\epsilon}$, and for any $r \in \{1, \ldots, n_j\}$,

$$\zeta^{(j,\epsilon)}_r = \int_{B_\epsilon(y^*_j)\times\mathbb{R}^+} \left(\prod_{i=1}^{r} \gamma_i(y)\right) v^r\, \mathrm{e}^{-v\, q^{\mathbf{0}}_{j,\epsilon}(y)}\, \rho(\mathrm{d}v|y) \frac{\alpha(\mathrm{d}y)}{\alpha(B_\epsilon(y^*_j))}.$$

Hence, one can again conclude that

$$V^{(n_j)}_{j,\epsilon} = \alpha(B_\epsilon(y^*_j))\, \zeta^{(n_j)}_j + o(\alpha(B_\epsilon(y^*_j)))$$

as $\epsilon \downarrow 0$, where $\zeta^{(n_j)}_j = \prod_{i\in C_j} \gamma_i(y^*_j) \int_{\mathbb{R}^+} v^{n_j}\, \mathrm{e}^{-v[f(y^*_j)+g_m(y^*_j)]} \rho(\mathrm{d}v|y^*_j)$, and for any $i$ in $C_j$, $\gamma_i(y^*_j) = K(x_i, y^*_j)\, \tau(\mathrm{d}x_i)$. This implies that

$$\mathbb{E}\left[\mathrm{e}^{-\mu(f)} \,\middle|\, \boldsymbol{X}, \boldsymbol{Y}\right] = \frac{\mathbb{E}\left[\mathrm{e}^{-\mu(f+g_m)}\right] \prod_{j=1}^{k} \int_{\mathbb{R}^+} \mathrm{e}^{-v\, f(y^*_j)}\, v^{n_j}\, \mathrm{e}^{-v g_m(y^*_j)} \rho(\mathrm{d}v|y^*_j)}{\mathbb{E}\left[\mathrm{e}^{-\mu(g_m)}\right] \prod_{j=1}^{k} \int_{\mathbb{R}^+} v^{n_j} \mathrm{e}^{-v g_m(y^*_j)} \rho(\mathrm{d}v|y^*_j)}$$

from which the statement in (i) follows.

One can analogously determine the conditional distribution of $\boldsymbol{Y}$, given $\boldsymbol{X}$ and $\boldsymbol{\Pi}_n$. Indeed, if $\boldsymbol{n} = (n_1, \ldots, n_k)$, one has

$$\begin{aligned} &P\left[\bigcap_{i=1}^{n} \{Y_i \in B_\epsilon(y^*_{j(i)})\} \cap \bigcap_{i=1}^{n} \{X_i \in B_\epsilon(x_i)\} \cap \{\boldsymbol{\Pi}_n = (k, \boldsymbol{n})\}\right] \\ &\quad = \mathbb{E}\left[\mathrm{e}^{-\mu(g_m)} \prod_{j=1}^{k} \prod_{i\in C_j} \int_{B_\epsilon(y^*_j)\times B_\epsilon(x_i)} K(x, y)\, \tau(\mathrm{d}x)\, \mu(\mathrm{d}y)\right] \\ &\quad = \mathrm{e}^{-\psi_n u(g_m \mathbb{I}_{\mathbb{Y}^*_\epsilon})} \prod_{j=1}^{k} \mathbb{E}\left[\mathrm{e}^{-\mu(g_m \mathbb{I}_{B_\epsilon(y^*_j)})} \prod_{i\in C_j} \int_{B_\epsilon(y^*_j)\times B_\epsilon(x_i)} K(x, y)\, \tau(\mathrm{d}x)\, \mu(\mathrm{d}y)\right]. \end{aligned} \tag{19}$$

Suppose again $C_j = \{1, \ldots, n_j\}$ and, after setting $h^{\boldsymbol\lambda}_{j,\epsilon} = [g_m + \sum_{i=1}^{n_j} \lambda_i\, \gamma_{i,\epsilon}]\, \mathbb{I}_{B_\epsilon(y^*_j)}$, one can proceed in a similar fashion as before to obtain the following representation for the probability in

Equation (19):

$$\mathrm{e}^{-\psi_v(g_m\,\mathbb{I}_{\mathbb{Y}^*_\epsilon})}\prod_{j=1}^{k}(-1)^{n_j}\,\frac{\partial^{n_j}}{\partial\lambda_1\ldots\partial\lambda_{n_j}}\mathbb{E}\left[\mathrm{e}^{-\mu(h^{\boldsymbol{\lambda}}_{j,\epsilon})}\right]\Big|_{\boldsymbol{\lambda}=\mathbf{0}}=\mathrm{e}^{-\psi_v(g_m)}\,\prod_{j=1}^{k}W^{(n_j)}_{j,\epsilon},$$

where $\mathbb{Y}^*_\epsilon=(\cup_{j=1}^k B_\epsilon(y_j^*))^c$,

$$\begin{aligned}W^{(n_j)}_{j,\epsilon}&=\mathrm{e}^{\psi_v(h^{\mathbf{0}}_{j,\epsilon})}\,(-1)^{n_j}\,\frac{\partial^{n_j}}{\partial\lambda_1\ldots\partial\lambda_{n_j}}\mathrm{e}^{-\psi_v(h^{\boldsymbol{\lambda}}_{j,\epsilon})}\Big|_{\boldsymbol{\lambda}=\mathbf{0}}\\&=\alpha(B_\epsilon(y_j^*))\sum_{i=0}^{n_j-1}\binom{n_j-1}{i}\zeta^{(j,\epsilon)}_{n_j-i}\,W^{(i)}_{j,\epsilon}\end{aligned}$$

and

$$\zeta^{(j,\epsilon)}_r=\int_{B_\epsilon(y_j^*)}\left(\prod_{i=1}^{r}\gamma_{i,\epsilon}(y)\right)\int_{\mathbb{R}^+}v^r\,\mathrm{e}^{-vg_m(y)}\,\rho(\mathrm{d}v|y)\,\frac{\alpha(\mathrm{d}y)}{\alpha(B_\epsilon(y_j^*))}.$$

Let $\delta$ denote the counting measure on $\{1,\ldots,n\}\times(\cup_{k=1}^n A_{k,n})$, so that the density of the vector $(\boldsymbol{Y},\boldsymbol{X},\boldsymbol{\Pi}_n)$ is absolutely continuous with respect to $\alpha^k\times\tau^n\times\delta$. The corresponding density evaluated at $(\boldsymbol{y}^*,\boldsymbol{x},\boldsymbol{\pi}_n)$ coincides with

$$\begin{aligned}&f(\boldsymbol{y}^*,\boldsymbol{x},\boldsymbol{\pi}_n)\\&\quad=\mathrm{e}^{-\psi_v(g_m)}\prod_{j=1}^{k}\left\{\prod_{i\in C_j}K(x_i,y_j^*)\right\}\int_{\mathbb{R}^+}v^{n_j}\mathrm{e}^{-vg_m(y_j^*)}\rho(\mathrm{d}v|y_j^*),\end{aligned}$$

where $\boldsymbol{\pi}_n=(k,n_1,\ldots,n_k)$. At this point, one can easily determine the distribution of $\boldsymbol{Y}$ given $(\boldsymbol{X},\boldsymbol{\Pi}_n)$, which admits density on $\mathbb{Y}^k$ (with respect to $\alpha^k$) coinciding with $f(\boldsymbol{y}^*\,|\,\boldsymbol{x},\boldsymbol{\pi}_n)\propto\prod_{j=1}^k\omega_{n_j}(y_j^*)\,\prod_{i\in C_j}K(x_i,y_j^*)$ from which Equations (16) and (17) follow. ■

## Acknowledgements

The authors are grateful to an associate editor and two anonymous referees for their valuable comments. A. Lijoi and I. Prünster are partially supported by MiUR, Grants 2006/ 134525 and 2006/133449, respectively.

## References

[1] L.F. James, *Poisson process partition calculus with applications to exchangeable models and Bayesian nonparametrics*, Manuscript, MatharXiv arXiv:math/0205093v1, 2002.
[2] A.Y. Lo, *On a class of Bayesian nonparametric estimates: i. density estimates*, Ann. Statist. 12 (1984), pp. 351–357.
[3] A.Y. Lo and C.-S. Weng, *On a class of Bayesian nonparametric estimates: II. Hazard rate estimates*, Ann. Inst. Statist. Math. 41 (1989), pp. 227–245.
[4] J. Pitman, *Combinatorial stochastic processes. Ecole d'Eté de Probabilités de Saint-Flour XXXII–2002*, Lect. Notes Math. 1875, Springer, New York, 2006.
[5] L.F. James, *Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages*, Ann. Statist. 33 (2005), pp. 1771–1799.
[6] O. Aalen, *Nonparametric inference for a family of counting processes*, Ann. Statist. 6 (1978), pp. 701–726.
[7] R.L. Dykstra and P.W. Laud, *A Bayesian nonparametric approach to reliability*, Ann. Statist. 9 (1981), pp. 356–367.
[8] H. Ishwaran and L.F. James, *Computational methods for multiplicative intensity models using weighted gamma processes: proportional hazards, marked point processes, and panel count data*, J. Amer. Stat. Assoc. 99 (2004), pp. 175–190.
[9] L.F. James, *Bayesian calculus for gamma processes with applications to semiparametric intensity models*, Sankhya 65 (2003), pp. 179–206.

[10] L.E. Nieto-Barajas and S.G. Walker, *Bayesian nonparametric survival analysis via Lévy driven Markov processes*, Statist. Sinica 14 (2004), pp. 1127–1146.

[11] R. Wolpert and K. Ickstadt, *Poisson/gamma random field models for spatial statistics*, Biometrika 85 (1998), pp. 251–267.

[12] L.F. James, *Poisson calculus for spatial neutral to the right processes*, Ann. Statist. 34 (2006), pp. 416–440.

[13] K. Doksum, *Tailfree and neutral random probabilities and their posterior distributions*, Ann. Probab. 2 (1974), pp. 183–201.

[14] I. Epifani, A. Lijoi, and I. Prünster, *Exponential functionals and means of neutral-to-the-right priors*, Biometrika 90 (2003), pp. 791–808.

[15] T.S. Ferguson and E.G. Phadia, *Bayesian nonparametric estimation based on censored data*, Ann. Statist. 7 (1979), pp. 163–186.

[16] S. Walker, and P. Damien, *A full Bayesian nonparametric analysis involving a neutral to the right process*, Scand. J. Statist. 25 (1998), pp. 669–680.

[17] S. Walker and P. Muliere, *Beta-stacy processes and a generalization of the pólya-urn scheme*, Ann. Statist. 25 (1997), pp. 1762–1780.

[18] N.L. Hjort, *Nonparametric Bayes estimators based on beta processes in models for life history data*, Ann. Statist. 18 (1990), pp. 1259–1294.

[19] P. De Blasi and N.L. Hjort, *Proportional hazard regression models. Bayesian analysis using Beta processes*, Scand. J. Statist. 34 (2007), pp. 229–257.

[20] Y. Kim, *Nonparametric Bayesian estimators for counting processes*, Ann. Statist. 27 (1999), pp. 562–588.

[21] Y. Kim and J. Lee, *Bayesian analysis of proportional hazard models*, Ann. Statist. 31 (2003), pp. 493–511.

[22] A. Lijoi and I. Prünster, *A note on the problem of heaps*, Sankhya 66 (2004), pp. 234–242.

[23] I. Prünster, *Random probability measures derived from increasing additive processes and their application to Bayesian statistics*, Ph.D. thesis, University of Pavia, 2002.

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

[25] L.F. James, A. Lijoi, and I. Prünster, *Posterior analysis for normalized random measures with independent increments, Scand. J. Statist.* (2008). DOI: 10.1111/j.1467-9469.2008.00609.x.

[26] G.M. Constantine and T.H. Savits, *A stochastic process interpretation of partition identities*, SIAM J. Discrete Math. 7 (1994), pp. 194–202.

[27] J.F.C. Kingman, *Completely random measures*, Pacific J. Math. 21 (1967), pp. 59–78.

[28] D.J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer, New York, 1988.

[29] J.F.C. Kingman, *Poisson Processes*, Oxford University Press, Oxford, 1993.

[30] L.F. James, *A simple proof of the almost sure discreteness of a class of random measures*, Statist. Probab. Lett. 65 (2003), pp. 363–368.