Adv. Appl. Prob. (SGSA) **43,** 301–307 (2011)
Printed in Northern Ireland

ON THE CONDITIONAL DISTRIBUTIONS OF SPATIAL POINT PROCESSES

FRANÇOIS CARON * AND
PIERRE DEL MORAL,* *Université Bordeaux 1*
ARNAUD DOUCET,** *University of British Columbia*
MICHELE PACE,* *Université Bordeaux 1*

Abstract

We consider the problem of estimating a latent point process, given the realization of another point process. We establish an expression for the conditional distribution of a latent Poisson point process given the observation process when the transformation from the latent process to the observed process includes displacement, thinning, and augmentation with extra points. Our original analysis is based on an elementary and self-contained random measure theoretic approach. This simplifies and complements previous derivations given in Mahler (2003), and Singh, Vo, Baddeley and Zuyev (2009).

Keywords: Filtering; multitarget tracking; spatial point process; probability hypothesis density filter

2010 Mathematics Subject Classification: Primary 62M30
Secondary 93E11; 60D05

1. Introduction

Spatial point processes occur in a wide variety of scientific disciplines including environmetrics, epidemiology, and seismology; see [1] and [7] for recent books on the subject. In this paper we are interested in scenarios where the spatial point process of interest is unobserved and we only have access to another spatial point process which is obtained from the original process through displacement, thinning, and augmentation with extra points. Such problems arise in forestry [2], [3], but our motivation for this work stems from target tracking applications [5], [6], [9]. In this context, we want to infer the number of targets and their locations; this number can vary as targets enter and exit the surveillance area. We only have access to measurements from a sensor. Some targets may not be detected by the sensor and additionally this sensor also provides us with a random number of false measurements.

From a mathematical point of view, we are interested in the computation of the conditional distributions of a sequence of random measures with respect to a sequence of noisy and partial observations given by spatial point processes. Recently, a few articles have addressed this problem. In a seminal paper [5], Mahler proposed an original and elegant multiobject filtering algorithm known as the PHD (probability hypothesis density) filter which relies on a first-order moment approximation of the posterior. The mathematical techniques used by Mahler are essentially based on random finite sets techniques including set derivatives and probability

Received 11 May 2010; revision received 4 February 2011.
* Postal address: INRIA Sud-Ouest and Institut de Mathématiques de Bordeaux, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence cedex, France.
** Postal address: Department of Statistics, 333-6356 Agricultural Road, Vancouver BC, V6T 1Z2, Canada.
Email address: arnaud@cs.ubc.ca

generating functionals. In a more recent article [6], Singh *et al.* clarified some important technicalities concerning the use of the derivatives of the joint probability generating functionals to characterise conditional distributions. They proposed a simplified derivation of the PHD filter and extended this algorithm to include second moment information. An alternative way to obtain such conditional distributions appeared in [4] and, using Janossy densities, in [8].

The main contribution of this paper is to propose an original analysis based on a self-contained random measure theoretic approach. The elementary techniques developed in this paper complement the more traditional random finite sets analysis involving symmetrization techniques or related to other technicalities associated with the computation of moment generating function derivatives.

The rest of this paper is organised as follows. In Section 2 we first present a static model associated to a pair of signal-observation Poisson point processes. We establish a functional representation of the conditional distribution of a Poisson signal process with respect to noisy and partial observations. The proof is elementary. It is extended in Section 3 to dynamic models in order to establish the PHD equations [5], [6]. We end this introductory section with some standard notation used in the paper.

We respectively denote by $\mathcal{M}(E)$, $\mathcal{P}(E)$, and $\mathcal{B}(E)$ the set of all finite positive measures on some measurable space $(E, \mathcal{E})$, the set of all probability measures, and the Banach space of all bounded and measurable real-valued functions. For $\mu \in \mathcal{M}(E)$ and $f \in \mathcal{B}(E)$, we let

$$\mu(f) = \int \mu(\mathrm{d}x) f(x)$$

be the Lebesgue integral. The Dirac measure at $a \in E$ is denoted δ_a. We also denote by $\mu^{\otimes p}$ the product measure of $\mu \in \mathcal{M}(E)$ on the product space E^p.

Let $G\colon x \in E \mapsto G(x) \in [0, \infty)$ be a bounded nonnegative potential function. Then $\Psi_G(\eta) \in \mathcal{P}(E)$ denotes the probability measure admitting a density $G(x)/\eta(G)$ with respect to a measure $\eta \in \mathcal{M}(E)$.

For every sequence of points $x = (x^i)_{1 \le i \le k}$ in E and every $0 \le p \le k$, we denote by $m_p(x)$ the occupation measure of the first p coordinates:

$$m_p(x) = \sum_{1 \le i \le p} \delta_{x^i}.$$

For $p = 0$, we use the convention that $m_0(x) = 0$, the null measure on E. We recall that a bounded and positive integral operator $f \mapsto L(f)$ from $\mathcal{B}(E_2)$ into $\mathcal{B}(E_1)$ is such that the functions

$$x \mapsto L(f)(x) = \int_{E_2} L(x, \mathrm{d}y) f(y)$$

are $\mathcal{E}_1$-measurable and bounded for some measures $L(x, \cdot) \in \mathcal{M}(E_2)$. These operators also generate a dual operator $\mu \mapsto \mu L$ from $\mathcal{M}(E_1)$ into $\mathcal{M}(E_2)$ defined by $(\mu L)(f) = \mu(L(f))$. A Markov kernel is obtained when $L(x, \cdot) \in \mathcal{P}(E)$ for any x.

In this paper we shall add an auxiliary 'death' state c to the state space E_1 and another auxiliary 'death' state d to the state space E_2. The functions $f \in \mathcal{B}(E_1)$ are extended to the augmented space $E_1 \cup \{c\}$ by setting $f(c) = 0$. Similarly, the functions $f \in \mathcal{B}(E_2)$ are extended to the augmented space $E_2 \cup \{d\}$ by setting $f(d) = 0$.

2. Conditional distributions for Poisson processes

Assume that the unobserved point process is a finite Poisson point process

$$\mathcal{X} = \sum_{1 \le i \le N} \delta_{X^i}$$

with intensity measure γ on some measurable state space $(E_1, \mathcal{E}_1)$. We set

$$\eta(\mathrm{d}x) = \frac{\gamma(\mathrm{d}x)}{\gamma(1)}.$$

The observed point process consists of a collection of random observations directly generated by a random number of points of $\mathcal{X}$ plus some random observations unrelated to $\mathcal{X}$.

To describe more precisely this observed point process, we let α be a measurable function from E_1 into $[0, 1]$ and we consider a Markov transition $L(x, \mathrm{d}y)$ from E_1 to E_2. Given a realization of $\mathcal{X}$, every random point $X^i = x$ generates, with probability $\alpha(x)$, an observation Y'^i on E_2 with distribution $L(x, \mathrm{d}y)$; otherwise it goes into a death state d. Hence, $\alpha(x)$ measures the 'detectability' degree of x. In other words, a given point x generates a random observation in $E_2' = E_2 \cup \{d\}$ with distribution

$$L_d(x, \mathrm{d}y) = \alpha(x)\, L(x, \mathrm{d}y) + (1 - \alpha(x))\delta_d(\mathrm{d}y). \tag{2.1}$$

The resulting point process is the random measure $\sum_{1 \le i \le N} \delta_{Y'^i}$ on the augmented state space E_2'.

In addition to this point process we also observe an additional, and independent of $\mathcal{X}$, Poisson point process $\sum_{1 \le i \le N_c} \delta_{Y_c'^i}$ with intensity measure ν on E_2; this is known as the clutter noise in multitarget tracking.

In other words, we obtain a process on E_2' given by the random measure

$$\mathcal{Y}' = \sum_{1 \le i \le N} \delta_{Y'^i} + \sum_{1 \le i \le N_c} \delta_{Y_c'^i}.$$

The state d being unobservable, the observed point process is the random measure $\mathcal{Y}$ on E_2 given by

$$\mathcal{Y} = \sum_{1 \le i \le N} \mathbf{1}_{E_2}(Y'^i)\delta_{Y'^i} + \sum_{1 \le i \le N_c} \delta_{Y_c'^i} = \mathcal{Y}' - N_d \delta_d = \sum_{1 \le i \le M} \delta_{Y^i},$$

where $N_d = (\sum_{1 \le i \le N} \mathbf{1}_d(Y'^i))$ corresponds to the number of undetected/dead points, and $M = N - N_d + N_c$ is the number of observed points.

Let $\mathcal{X}' = \mathcal{X} + N_c \delta_c$ be defined on $E_1' = E_1 \cup \{c\}$, where c is some cemetery state associated with clutter observations. We present in the following proposition an explicit integral representation of a version of the conditional distributions of $\mathcal{Y}'$ given $\mathcal{X}$ and $\mathcal{X}'$ given $\mathcal{Y}$.

Proposition 2.1. *A version of the conditional distribution of $\mathcal{Y}'$ given $\mathcal{X}$ is given, for any function $F \in \mathcal{B}(\mathcal{M}(E_2'))$, by*

$$\mathrm{E}(F(\mathcal{Y}') \mid \mathcal{X}) = \mathrm{e}^{-\nu(1)} \sum_{k \ge 0} \frac{1}{k!} \int_{(E_2')^{k+N}} F(m_k(y_c') + m_N(y'))\nu^{\otimes k}(\mathrm{d}y_c') \prod_{i=1}^{N} L_d(X^i, \mathrm{d}y'^i). \tag{2.2}$$

We further assume that $\nu \ll \lambda$ *and* $L(x, \cdot) \ll \lambda$ *for any* $x \in E_1$ *and some reference measure* $\lambda \in \mathcal{M}(E_2)$*, with Radon Nikodym derivatives given by*

$$g(x, y) = \frac{\mathrm{d}L(x, \cdot)}{\mathrm{d}\lambda}(y) \quad \text{and} \quad h(y) = \frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(y), \tag{2.3}$$

and such that $h(y) + \gamma(\alpha g(\cdot, y)) > 0$ *for any* $y \in E_2$*.*

In this situation, a version of the conditional distribution of $\mathcal{X}'$ *given the observation point process* $\mathcal{Y}$ *is given, for any function* $F \in \mathcal{B}(\mathcal{M}(E_1'))$*, by*

$$\mathrm{E}(F(\mathcal{X}') \mid \mathcal{Y}) = \mathrm{e}^{-\gamma(1-\alpha)} \sum_{k \geq 0} \frac{\gamma(1-\alpha)^k}{k!} \int_{(E_1')^{k+M}} F(m_k(x') + m_M(x)) \Psi_{(1-\alpha)}(\eta)^{\otimes k}(\mathrm{d}x') \times \prod_{i=1}^{M} Q(Y^i, \mathrm{d}x^i), \tag{2.4}$$

where Q *is a Markov transition from* E_2 *into* E_1' *defined by*

$$Q(y, \mathrm{d}x) = (1 - \beta(y)) \Psi_{\alpha g(\cdot, y)}(\eta)(\mathrm{d}x) + \beta(y) \delta_c(\mathrm{d}x)$$

with

$$\beta(y) = \frac{h(y)}{h(y) + \gamma(\alpha g(\cdot, y))}. \tag{2.5}$$

Proof. The proof of the first assertion in (2.2) is elementary and is thus omitted. We provide here a proof of the second result given in (2.4). First, we observe that the random measure

$$\mathcal{Z} = \sum_{1 \leq i \leq N} \delta_{(X^i, Y'^i)} + \sum_{1 \leq i \leq N_c} \delta_{(c, Y_c'^i)} = \sum_{1 \leq i \leq N + N_c} \delta_{(Z_1^i, Z_2^i)}$$

is a Poisson point process in $E' = E_1' \times E_2'$. More precisely, the random variable $N + N_c$ is a Poisson random variable with parameter $\kappa = \gamma(1) + \nu(1)$, and $(Z_1^i, Z_2^i)_{i \geq 0}$ is a sequence of independent random variables with common distribution

$$\Gamma(\mathrm{d}(z_1, z_2)) = \eta'(\mathrm{d}z_1) L'(z_1, \mathrm{d}z_2) \quad \text{with} \quad \kappa \eta' = \gamma(1)\eta + \nu(1)\delta_c,$$

$$L'(z_1, \mathrm{d}z_2) = \mathbf{1}_{E_1}(z_1)\, L_d(z_1, \mathrm{d}z_2) + \mathbf{1}_c(z_1) \bar{\nu}(\mathrm{d}z_2) \quad \text{with} \quad \bar{\nu}(\mathrm{d}z_2) = \frac{\nu(\mathrm{d}z_2)}{\nu(1)}.$$

From the joint distribution $\Gamma(\mathrm{d}(z_1, z_2))$, we can obtain the conditional distribution $L'_{\eta'}(z_2, \mathrm{d}z_1)$ of Z_1 given $Z_2 = z_2$ using the easily checked reversal formula, i.e. the Bayes rule

$$\eta'(\mathrm{d}z_1) L'(z_1, \mathrm{d}z_2) = (\eta' L')(\mathrm{d}z_2)\, L'_{\eta'}(z_2, \mathrm{d}z_1).$$

This yields

$$L'_{\eta'}(z_2, \mathrm{d}z_1) = \mathbf{1}_d(z_2) \Psi_{(1-\alpha)}(\eta)(\mathrm{d}z_1) + \mathbf{1}_{E_2}(z_2)\, Q(z_2, \mathrm{d}z_1).$$

Hence, we can conclude that, for any function $F \in \mathcal{B}(\mathcal{M}(E_1'))$,

$$\mathrm{E}(F(\mathcal{Z}_1) \mid \mathcal{Z}_2) = \int_{(E_1')^{N+N_c}} F(m_{N+N_c}(z_1)) \prod_{i=1}^{N+N_c} L'_{\eta'}(Z_2^i, \mathrm{d}z_1^i),$$

where $\mathcal{Z}_j$ stands for the jth marginal of $\mathcal{Z}$, with $j \in \{1, 2\}$. The end of the proof is now a direct consequence of the facts that $(\mathcal{Z}_1, \mathcal{Z}_2) = (\mathcal{X}', \mathcal{Y}')$, $\mathrm{E}(F(\mathcal{X}') \mid \mathcal{Y}) = \mathrm{E}(\mathrm{E}(F(\mathcal{X}') \mid \mathcal{Y}') \mid \mathcal{Y})$, and

$$\mathrm{E}(F(\mathcal{Y}') \mid \mathcal{Y}) = \mathrm{e}^{-\gamma(1-\alpha)} \sum_{k\geq 0} \frac{\gamma(1-\alpha)^k}{k!} F(k\delta_d + \mathcal{Y}),$$

for any function $F \in \mathcal{B}(\mathcal{M}(E_2'))$ as N_d follows a Poisson distribution of parameter $\gamma(1-\alpha)$. This completes the proof of the proposition.

The expressions of the conditional expectations of linear functionals of the random point processes $\mathcal{X}'$ and $\mathcal{X}$ given the point process $\mathcal{Y}$ follow straightforwardly from the previous proposition. Recall that $f(c) = 0$ by convention.

Corollary 2.1. *For any function $f \in \mathcal{B}(E_1')$, we have*

$$\begin{aligned}\mathrm{E}(\mathcal{X}'(f) \mid \mathcal{Y}) &= \mathrm{E}(\mathcal{X}(f) \mid \mathcal{Y}) \\ &= \mathrm{e}^{-\gamma(1-\alpha)} \sum_{k\geq 0} \frac{\gamma(1-\alpha)^k}{k!} \bigg(k\, \Psi_{(1-\alpha)}(\eta)(f) + \int \mathcal{Y}(\mathrm{d}y) Q(f)(y) \bigg) \\ &= \gamma((1-\alpha)f) + \int \mathcal{Y}(\mathrm{d}y)(1-\beta(y))\Psi_{\alpha g(\cdot,y)}(\eta)(f).\end{aligned}$$

In particular, the conditional expectation of the number of points N in $\mathcal{X}$ given the observations is given by

$$\mathrm{E}(N \mid \mathcal{Y}) = \mathrm{E}(\mathcal{X}(1) \mid \mathcal{Y}) = \gamma(1-\alpha) + \mathcal{Y}(1-\beta).$$

3. Spatial filtering models and probability hypothesis density equations

We show here how the results obtained in Proposition 2.1 and Corollary 2.1 allow us to establish directly the PHD filter equations [5], [6].

In what follows the parameter n is interpreted as a discrete-time index. We consider a collection of measures $\mu_n \in \mathcal{M}(E_1)$ and a collection of positive operators R_{n+1} from E_1 into E_1.

We then define recursively a sequence of random measures $\mathcal{X}_n$ and $\mathcal{Y}_n$ on E_1 and E_2 as follows. The initial measure $\mathcal{X}_0$ is a Poisson point process with intensity measure $\gamma_0 = \mu_0$ on E_1. Given a realization of $\mathcal{X}_0$, the corresponding observation process $\mathcal{Y}_0$ on E_2 is defined as in Section 2 with a detection function α_0 on E_1, a clutter intensity measure ν_0, and some Markov transitions $L_{d,0}$ and L_0 defined as in (2.1) and satisfying (2.3) for some reference measure λ_0, and some functions h_0 and g_0. From Corollary 2.1 we have, for any function $f \in \mathcal{B}(E_1)$,

$$\begin{aligned}\hat{\gamma}_0(f) &= \mathrm{E}(\mathcal{X}_0(f) \mid \mathcal{Y}_0) \\ &= \gamma_0((1-\alpha_0)f) + \int \mathcal{Y}_0(\mathrm{d}y)(1-\beta_0(y))\Psi_{\alpha_0 g_0(\cdot,y)}(\gamma_0)(f),\end{aligned}$$

where the function β_0 is defined as in (2.5), by substituting (α_0, h_0, g_0) for (α, h, g). Given a realization of the pair random sequences $(\mathcal{X}_p, \mathcal{Y}_p)$, with $0 \leq p \leq n$, the pair of random measures $(\mathcal{X}_{n+1}, \mathcal{Y}_{n+1})$ is *defined* as follows. We set $\mathcal{X}_{n+1}$ to be a Poisson point process with intensity measure γ_{n+1} defined by the following recursions for any function $f \in \mathcal{B}(E_1)$:

$$\begin{aligned}\hat{\gamma}_n(f) &= \gamma_n((1-\alpha_n)f) + \int \mathcal{Y}_n(\mathrm{d}y)(1-\beta_n(y))\Psi_{\alpha_n g_n(\cdot,y)}(\gamma_n)(f), \\ \gamma_{n+1} &= \hat{\gamma}_n R_{n+1} + \mu_{n+1}.\end{aligned}$$

In the context of spatial branching processes, μ_n stands for the intensity measure of a spontaneous birth model while R_{n+1} represents the first moment transport kernel associated with a spatial branching type mechanism. For example, assume that each point $X_n^i = x$ at time n dies with probability $\rho(x)$ or survives and evolves according to a Markov kernel K_{n+1} from E_1 into E_1. Then R_{n+1} corresponds to

$$R_{n+1}(x, \mathrm{d}x') = (1 - \rho(x))K_{n+1}(x, \mathrm{d}x').$$

It is also possible to modify R_{n+1} to include some spawning points [5], [6], [9]. In addition, given a realization of $\mathcal{X}_{n+1}$, the corresponding observation process $\mathcal{Y}_{n+1}$ is defined as in Section 2 with a detection function α_{n+1} on E_1, a clutter intensity measure ν_{n+1}, and some Markov transitions $L_{d,(n+1)}$ and L_{n+1} defined as in (2.1) and satisfying (2.3) for some reference measure λ_{n+1}, and some functions h_{n+1} and g_{n+1}. We let $N_{c,n}$ be the number of death states c associated with clutter observations at time n, and let M_n be the number of observations at time n.

The following elementary corollary proves that the PHD filter propagates the first moment of the multitarget posterior distribution of the filtering model defined above. This is a direct consequence of Proposition 2.1 and Corollary 2.1.

Corollary 3.1. *An integral version of the conditional distribution of* $\mathcal{X}_n' = \mathcal{X}_n + N_{c,n}\delta_c$ *given the filtration* $\mathcal{F}_n^Y = \sigma(\mathcal{Y}_p,\ 0 \le p \le n)$ *generated by the observation point processes*

$$\mathcal{Y}_p = \sum_{1 \le i \le M_p} \delta_{Y_p^i},$$

from the origin $p = 0$ *up to the current time* $p = n$*, is given, for any function* $F \in \mathcal{B}(\mathcal{M}(E_1'))$*, by*

$$\begin{aligned} \mathrm{E}(F(\mathcal{X}_n') \mid \mathcal{F}_n^Y) &= \mathrm{e}^{-\gamma_n(1-\alpha_n)} \sum_{k \ge 0} \frac{\gamma_n(1-\alpha_n)^k}{k!} \int_{(E_1')^{k+M_n}} F(m_k(x') + m_{M_n}(x))\Psi_{(1-\alpha_n)}(\gamma_n)^{\otimes k}(\mathrm{d}x') \\ &\qquad \times \prod_{i=1}^{M_n} Q_n(Y_n^i, \mathrm{d}x^i) \end{aligned}$$

with the Markov transitions

$$Q_n(y, \mathrm{d}x) = (1 - \beta_n(y))\Psi_{\alpha_n g_n(\cdot, y)}(\gamma_n)(\mathrm{d}x) + \beta_n(y)\delta_c(\mathrm{d}x).$$

In particular, the random measures γ_n *and* $\hat{\gamma}_n$ *defined below coincide with the first moment of the random measures* $\mathcal{X}^n$ *given the sigma-fields* $\mathcal{F}_{n-1}^Y$ *and* $\mathcal{F}_n^Y$*; that is, for any function* $f \in \mathcal{B}(E_1)$*, we have*

$$\gamma_n(f) = \mathrm{E}(\mathcal{X}_n(f) \mid \mathcal{F}_{n-1}^Y) \quad \textit{and} \quad \hat{\gamma}_n(f) = \mathrm{E}(\mathcal{X}_n(f) \mid \mathcal{F}_n^Y).$$

References

[1] Daley, D. J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*, Springer, New York.
[2] Lund, J. and Rudemo, M. (2000). Models for point processes observed with noise. *Biometrika*, **87,** 235–249.
[3] Lund, J. and Thönnes, E. (2004). Perfect simulation and inference for point processes given noisy observations. *Comput. Statist.* **19,** 317–336.
[4] Lund, J., Penttinen, A. and Rudemo, M. (1999). Bayesian analysis of spatial point patterns from noisy observations. Tech. Rep., Department of Mathematics and Physics, The Royal Veterinary and Agricultural University.

[5] Mahler, R. P. S. (2003). Multi-target Bayes filtering via first-order multitarget moments. *IEEE Trans. Aerospace Electronic Systems* **39,** 1152–1178.

[6] Singh, S. S., Vo, B.-N., Baddeley, A. and Zuyev, S. (2009). Filters for spatial point processes. *SIAM J. Control Optimization* **48,** 2275–2295.

[7] Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications*, 2nd edn. John Wiley, Chichester.

[8] Van Lieshout, M. N. M. and Baddeley, A. J. (2002). Extrapolating and interpolating spatial patterns. In *Spatial Cluster Modelling*, Chapman & Hall/CRC, Boca Raton, FL.

[9] Vo, B.-T., Vo, B.-N. and Cantoni, A. (2007). Analytic implementations of the cardinalized probability hypothesis density filter. *IEEE Trans. Signal Process.* **55,** 3553–3567.