Applied Probability Trust (4 February 2011)

ON THE CONDITIONAL DISTRIBUTIONS OF SPATIAL POINT PROCESSES

FRANÇOIS CARON,* *INRIA Sud-Ouest & Institut de Mathématiques de Bordeaux*

PIERRE DEL MORAL,** *INRIA Sud-Ouest & Institut de Mathématiques de Bordeaux*

ARNAUD DOUCET,*** *University of British Columbia*

MICHELE PACE,**** *INRIA Sud-Ouest & Institut de Mathématiques de Bordeaux*

Abstract

We consider the problem of estimating a latent point process, given the realization of another point process. We establish an expression of 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 [5], [6].

Keywords: filtering; multitarget tracking; spatial point processes; probability hypothesis density filter

2000 Mathematics Subject Classification: Primary 62M30

Secondary 93E11;60D05

* Postal address: Université Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France
** Postal address: Université Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France
*** Postal address: Department of Statistics, 333-6356 Agricultural Road, Vancouver BC, V6T 1Z2, Canada
**** Postal address: Université Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France

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 [3], [4] 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], R. Mahler has proposed an original and elegant multi-object 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 R. Mahler are essentially based on random finite sets techniques including set derivatives and probability generating functionals. In a more recent article [6], S.S. Singh, B.N. Vo, A. Baddeley and S. Zuyev have clarified some important technicalities concerning the use of the derivatives of the joint probability generating functionals to characterize conditional distributions. They have proposed a simplified derivation of the PHD filter and have extended this algorithm to include second moment information. An alternative way to obtain such conditional distributions appeared in [2] and, using Janossy densities, in [8].

The main contribution of this article 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 functions derivatives.

The rest of this article is organized 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 w.r.t. 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 notations used in the paper.

We denote respectively 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 $\mu(f) = \int \mu(dx)\ f(x)$ 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 $G : x \in E \mapsto G(x) \in [0, \infty)$ be a bounded non-negative 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 $x = (x^i)_{1 \leq i \leq k}$ in E and every $0 \leq p \leq k$, we denote by $m_p(x)$ the occupation measure of the first p coordinates $m_p(x) = \sum_{1 \leq i \leq p} \delta_{x^i}$. For $p = 0$, we use the convention $m_0(x) = 0$, 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

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

are $\mathcal{E}_1$-measurable and bounded for some measures $L(x, .) \in \mathcal{M}(E_2)$. 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 article, 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 the unobserved point process is a finite Poisson point process $\mathcal{X} = \sum_{1\leq i\leq N} \delta_{X^i}$ with intensity measure γ on some measurable state space $(E_1, \mathcal{E}_1)$. We set $\eta(dx) = \gamma(dx)/\gamma(1)$. 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 $L(x, dy)$ 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 $L(x, dy)$; 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

$$L_d(x, dy) = \alpha(x)\ L(x, dy) + (1-\alpha(x))\ \delta_d(dy). \tag{1}$$

The resulting point process is the random measure $\sum_{1\leq i\leq N} \delta_{Y'^i}$ 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 $\sum_{1\leq i\leq N_c} \delta_{Y_c'^i}$ 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

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

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

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

where $N_d = \left(\sum_{1\leq i\leq N} 1_d(Y'^i)\ \right)$ 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 $F \in \mathcal{B}\left(\mathcal{M}(E_2')\right)$ by*

$$\mathbb{E}\left(F\left(\mathcal{Y}'\right)|\mathcal{X}\right) = e^{-\nu(1)} \sum_{k\geq 0} \frac{1}{k!} \int_{\left(E_2'\right)^{k+N}} F\left(m_k(y_c') + m_N(y')\right)\ \nu^{\otimes k}(dy_c')\ \prod_{i=1}^{N} L_d(X^i, dy'^i) \tag{2}$$

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

$$g(x,y) = \frac{dL(x, \centerdot)}{d\lambda}(y) \quad \text{and} \quad h(y) = \frac{d\nu}{d\lambda}(y) \tag{3}$$

and such that $h(y) + \gamma(\alpha g(\centerdot, y)) > 0$, 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 $F \in \mathcal{B}\left(\mathcal{M}(E_1')\right)$ by

$$\mathbb{E}\left(F\left(\mathcal{X}'\right)|\mathcal{Y}\right)$$

$$= e^{-\gamma(1-\alpha)} \sum_{k\geq 0} \frac{\gamma(1-\alpha)^k}{k!} \int_{\left(E_1'\right)^{k+M}} F\left(m_k(x') + m_M(x)\right)\ \Psi_{(1-\alpha)}(\eta)^{\otimes k}\left(dx'\right)\ \prod_{i=1}^{M} Q\left(Y^i, dx^i\right) \tag{4}$$

where Q is a Markov transition from E_2 into E_1' defined by the following formula

$$Q(y,dx) = (1-\beta(y))\ \ \Psi_{\alpha g(\centerdot,y)}(\eta)(dx) + \beta(y)\ \delta_c(dx) \tag{5}$$

with

$$\beta(y) = \frac{h(y)}{h(y) + \gamma(\alpha g(\centerdot, y))}\ . \tag{6}$$

Proof:

The proof of the first assertion in Eq. (2) is elementary, thus it is skipped. We provide here a proof of the second result given in Eq. (4). First, we observe that the random measure

$$\mathcal{Z} = \sum_{1\leq i\leq N} \delta_{(X^i, Y'^i)} + \sum_{1\leq i\leq N_c} \delta_{(c, Y_c'^i)} = \sum_{1\leq i\leq N+N_c} \delta_{(Z_1^i, Z_2^i)} \tag{7}$$

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

$$\Gamma(d(z_1,z_2)) = \eta'(dz_1)L'(z_1,dz_2) \qquad \text{with} \qquad \kappa\eta' = \gamma(1)\ \eta + \nu(1)\ \delta_c\ ,$$
$$L'(z_1,dz_2) = 1_{E_1}(z_1)\ L_d(z_1,dz_2) + 1_c(z_1)\ \overline{\nu}(dz_2) \quad \text{with} \quad \overline{\nu}(dz_2) = \nu(dz_2)/\nu(1)\ .$$

From the joint distribution $\Gamma(d(z_1,z_2))$, we can obtain the conditional distribution $L'_{\eta'}(z_2,dz_1)$ of Z_1 given $Z_2 = z_2$ using the easily checked reversal formula, i.e. the Bayes rule

$$\eta'(dz_1)L'(z_1,dz_2) = (\eta' L')\,(dz_2)\ L'_{\eta'}(z_2,dz_1).$$

This yields

$$L'_{\eta'}(z_2,dz_1) = 1_d(z_2)\ \Psi_{(1-\alpha)}(\eta)(dz_1) + 1_{E_2}(z_2)\ Q(z_2,dz_1).$$

Hence we can conclude that for any function $F \in \mathcal{B}(\mathcal{M}(E'_1))$

$$\mathbb{E}\left(F(\mathcal{Z}_1)\,|\mathcal{Z}_2\right) = \int_{\left(E'_1\right)^{N+N_c}} F\left(m_{N+N_c}(z_1)\right)\ \prod_{i=1}^{N+N_c} L'_{\eta'}\left(Z_2^i, dz_1^i\right)$$

where $\mathcal{Z}_j$ stands for the j-th marginal of $\mathcal{Z}$, with $j \in \{1,2\}$. The end of the proof is now a direct consequence of the fact that $(\mathcal{Z}_1,\mathcal{Z}_2) = (\mathcal{X}',\mathcal{Y}')$, $\mathbb{E}\left(F\left(\mathcal{X}'\right)|\mathcal{Y}\right) = \mathbb{E}\left(\mathbb{E}\left(F(\mathcal{X}')\,|\mathcal{Y}'\right)|\mathcal{Y}\right)$ and

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

for any function $F \in \mathcal{B}(\mathcal{M}(E'_2))$ as N_d follows a Poisson distribution of parameter $\gamma(1-\alpha)$. This ends 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*

$$\begin{aligned}
\mathbb{E}\left(\mathcal{X}'(f)\mid \mathcal{Y}\right) &= \mathbb{E}\left(\mathcal{X}(f)\mid \mathcal{Y}\right) \\
&= e^{-\gamma(1-\alpha)} \sum_{k\geq 0} \frac{\gamma(1-\alpha)^k}{k!} \left(k\ \Psi_{(1-\alpha)}(\eta)\,(f) + \int \mathcal{Y}(dy) Q\,(f)\,(y)\right) \\
&= \gamma((1-\alpha)f) + \int\ \mathcal{Y}(dy)\,(1-\beta(y))\ \ \Psi_{\alpha g(\,\bullet\,,y)}(\eta)(f)\ . \qquad (8)
\end{aligned}$$

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

$$\mathbb{E}\left(N|\mathcal{Y}\right) = \mathbb{E}\left(\mathcal{X}(1) \mid \mathcal{Y}\right) = \ \gamma(1-\alpha) + \mathcal{Y}\left(1-\beta\right). \tag{9}$$

3. Spatial filtering models and probability hypothesis density equations

We show here how the results obtained in proposition 2.1 and corollary 2.1 allows 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 (1) and satisfying (3) for some reference measure λ_0 and some functions h_0 and g_0. From corollary 2.1, we have for any function $f \in \mathcal{B}(E_1)$

$$\begin{aligned}\widehat{\gamma}_0(f) =& \mathbb{E}\left(\mathcal{X}_0(f) \mid \mathcal{Y}_0\right)\\ &= \gamma_0((1-\alpha_0)f) + \int \ \mathcal{Y}_0(dy)\left(1-\beta_0(y)\right) \ \Psi_{\alpha_0 g_0(\,\cdot\,,y)}(\gamma_0)(f)\end{aligned}$$

with a function β_0 defined as in Eq. (6) by substituting (α_0, h_0, g_0) to (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 $f \in \mathcal{B}(E_1)$

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

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

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

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 (1) and satisfying (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 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 multi-target 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* $\mathcal{F}^Y_n = \sigma(\mathcal{Y}_p,\ 0 \le p \le n)$ *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 the following formula*

$$\mathbb{E}\left(F\left(\mathcal{X}'_n\right) \middle| \mathcal{F}^Y_n\right) = e^{-\gamma_n(1-\alpha_n)} \sum_{k\ge 0} \frac{\gamma_n(1-\alpha_n)^k}{k!}$$
$$\int_{(E'_1)^{k+M_n}} F\left(m_k(x') + m_{M_n}(x)\right)\ \Psi_{(1-\alpha_n)}(\gamma_n)^{\otimes k}(dx')\ \prod_{i=1}^{M_n} Q_n\left(Y_n^i, dx^i\right)$$

with the Markov transitions

$$Q_n(y, dx) = (1 - \beta_n(y))\ \Psi_{\alpha_n g_n(\bullet, y)}(\gamma_n)(dx) + \beta_n(y)\ \delta_c(dx)\ .$$

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

$$\gamma_n(f) = \mathbb{E}\left(\mathcal{X}_n(f) \mid \mathcal{F}^Y_{n-1}\right) \quad \text{and} \quad \widehat{\gamma}_n(f) = \mathbb{E}\left(\mathcal{X}_n(f) \mid \mathcal{F}^Y_n\right).$$

References

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

[2] Lund, J., Penttinen, A. and Rudemo, M. (1999). Bayesian analysis of spatial point patterns from noisy observations. *Technical report*. Department of Mathematics and Physics, The Royal Veterinary and Agricultural University.

[3] Lund, J. and Rudemo, M. (2000). Models for point processes observed with noise. *Biometrika* **87,** 235–249.

[4] Lund, J. and Thonnes, E. (2004). Perfect simulation and inference for point processes given noisy observations. *Computational Statistics* **19,** 317–336.

[5] Mahler, R. (2003). Multi-target Bayes filtering via first-order multi-target moments. *IEEE Transactions on Aerospace and Electronic Systems* **39,** 1152–1178.

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

[7] Stoyan, D. and Kendall, W. (1995). *Stochastic Geometry and Its Applications.* Wiley, New York.

[8] Van Lieshout, M. and Baddeley, A. (2002). Extrapolating and interpolating spatial patterns. In *Spatial cluster modelling.* Chapman & Hall/CRC.

[9] Vo, B., Vo, B. and Cantoni, A. (2007). Analytic implementations of the cardinalized probability hypothesis density filter. *IEEE Transactions on Signal Processing* **55,** 3553–3567.