# RANDOM FINITE SETS AND SEQUENTIAL MONTE CARLO METHODS IN MULTI-TARGET TRACKING

*Ba-Ngu Vo, Sumeetpal Singh and Arnaud Doucet\**

Department of Electrical and Electronic Engineering,
The University of Melbourne, Vic. 3010, Australia.
Department of Engineering,
Cambridge University, CB2 1PZ Cambridge, UK*
Email: bv@ee.mu.oz.au, ssss@ee.mu.oz.au, ad2@eng.cam.ac.uk

## ABSTRACT

Random finite set provides a rigorous foundation for optimal Bayes multi-target filtering. The major hurdle faced in Bayes multi-target filtering is the inherent computational intractability. Even the Probability Hypothesis Density (PHD) filter, which propagates only the first moment (or PHD) instead of the full multi-target posterior, still involves multiple integrals with no closed forms. In this paper, we highlight the relationship between Radon-Nikodym derivative and set derivative of random finite sets that enables a Sequential Monte Carlo (SMC) implementation of the optimal multi-target filter. In addition, a generalised SMC method to implement the PHD filter is also presented. The SMC PHD filter has an attractive feature-its computational complexity is independent of the (time-varying) number of targets.

*Keywords*: Multi-target Tracking, Optimal Filtering, Particle Methods, Point Processes, Random Sets, Sequential Monte Carlo.

## 1. INTRODUCTION

Mahler's Finite Set Statistics (FISST) [8], [9] provides a systematic treatment of multi-sensor multi-target tracking using random set theory. In FISST, the targets are treated as a single meta target, called the multi-target state, and the observations are treated as a single meta measurement, called the multi-target measurement. The multi-target state and multi-target measurement are modeled by random finite sets whose statistical behaviour are characterised by belief functions. The key to rigorously casting the multi-target estimation problem in a Bayesian framework is the notion of belief density, based on set-derivatives and set-integrals.

Analogous to the single-target case, Bayes multi-target tracking propagates the multi-target posterior density recursively in time [9]. This involves the evaluation of multiple set-integrals and the computational intractability is far more severe than its single-target counterpart. The Prob-ability Hypothesis Density (PHD) filter [10], which only propagates the 1st moment (or PHD) of the multi-target posterior, requires much less computational power. Unfortunately, this still involves multiple integrals that in gereral have no closed form.

Sequential Monte Carlo (SMC) methods are powerful tools which have had an impact on optimal (Bayesian) filtering [3]. The key idea in Monte Carlo methods is the approximation of integrals by random samples from the distribution of interest. To apply SMC methods to Bayes multi-target filtering in a principled way, we need to address firstly, the meaning of sampling from a belief density, and secondly, how to approximate a belief density from these samples. In this paper, we highlight the relationship between Radon-Nikodym derivative and set-derivative of random finite sets that provides answers to these questions. Consequently, a generic SMC implementation of the optimal Bayes multi-target filter can be proposed. In addition, a novel SMC implementation of the PHD filter which offers a much cheaper and more practical alternative is also presented. Both algorithms are general enough to capture non-linear non-Gaussian dynamics.

Section 2 of the paper briefly reviews the basics of random finite set, optimal Bayes multi-target tracking and describes the PHD or 1st moment filter. Section 3 presents SMC implementations of the Bayes multi-target filter and the PHD filter. Simulation results are presented in Section 4 and some concluding remarks are given in Section 5.

## 2. RANDOM FINITE SET AND BAYES MULTI-TARGET FILTERING

Multi-target state and multi-target measurement at time $k$ are naturally represented as finite sets $X_k$ and $Z_k$. For example, if at time $k$ there are $M(k)$ targets $x_{k,1}, \ldots, x_{k,M(k)}$ in the state space $E_s$ then, $X_k = \{x_{k,1}, \ldots, x_{k,M(k)}\} \subseteq E_s$. Similarly, if $N(k)$ observations $z_{k,1}, \ldots, z_{k,N(k)}$ in the

observation space $E_o$ are received at time $k$, then $Z_k = \{z_{k,1}, \ldots, z_{k,N(k)}\} \subseteq E_o$ where some of the $N(k)$ observations may be due to clutter. Analogous to single-target system, where uncertainty is characterised by modelling the state and measurement by random vectors, uncertainty in a multi-target system is characterised by modelling multi-target state and multi-target measurement as random finite sets (RFS) $\Xi_k$ and $\Sigma_k$ respectively. A formal definition of a RFS is given in Section 2.1.

Given a realisation $X_{k-1}$ of $\Xi_{k-1}$, the multi-target state at time $k$ can be modelled by the RFS

$$\Xi_k = S_k(X_{k-1}) \cup N_k(X_{k-1}) \qquad (1)$$

where $S_k(X_{k-1})$ denotes the RFS of targets that have survived at time $k$, $N_k(X_{k-1}) = B_k(X_{k-1}) \cup \Gamma_k$, $B_k(X_{k-1})$ is the RFS of targets spawned from $X_{k-1}$ and $\Gamma_k$ is the RFS of targets that appear spontaneously at time $k$. The statistical behaviour of the RFS $\Xi_k$ is characterised by the conditional probability "density" $f_{k|k-1}(X_k|X_{k-1})$ in an analogous fashion to the Markov transition density for random vector. The notion of probability density for RFS is formalised in Section 2.1.

Similarly, given a realisation $X_k$ of $\Xi_k$, the multi-target measurement can be modelled by the RFS

$$\Sigma_k = \Theta_k(X_k) \cup C_k(X_k) \qquad (2)$$

where $\Theta_k(X_k)$ denotes the RFS of measurements generated by $X_k$, and $C_k(X_k)$ denotes the RFS of clutter. The statistical behaviour of the RFS $\Sigma_k$ is described by the conditional probability "density" $g_k(Z_k|X_k)$ in an analogous fashion to the likelihood function for random vector observation.

Let $p_{k|k}(X_k|Z_{0:k})$ denote the multi-target posterior "density". Then, the optimal multi-target Bayes filter is given by the recursion

$$p_{k|k-1}(X_k|Z_{0:k-1})$$
$$= \int f_{k|k-1}(X_k|X) p_{k-1|k-1}(X|Z_{0:k-1}) \mu_s(dX) \quad (3)$$
$$p_{k|k}(X_k|Z_{0:k})$$
$$= \frac{g_k(Z_k|X_k) p_{k|k-1}(X_k|Z_{0:k-1})}{\int g_k(Z_k|X) p_{k|k-1}(X|Z_{0:k-1}) \mu_s(dX)}. \qquad (4)$$

where $\mu_s$ is a dominating measure to be discussed later in Section 2.1. The main difference between the recursion (3-4) and standard clutter-free single-target filtering is that $X_k$ and $Z_k$ can change dimension as $k$ changes.

## 2.1. Random Finite Sets

Given a closed and bounded subset $E$ of $\mathbf{R}^n$, let $\mathcal{F}(E)$ denote the collection of finite subsets of $E$. A random finite set (RFS) $\Xi$ on $E$ is defined as a measurable mapping

$$\Xi : \Omega \to \mathcal{F}(E).$$

The probability measure $P$ on the sample space $\Omega$ induces a probability law for $\Xi$, which can be specified in terms of probability distribution or belief function. The probability distribution $P_\Xi$ is defined for any subset[1] $\mathcal{T}$ of $\mathcal{F}(E)$ by

$$P_\Xi(\mathcal{T}) = P\Xi^{-1}(\mathcal{T}) = P(\{\omega : \Xi(\omega) \in \mathcal{T}\}).$$

The probability law for $\Xi$ can also be given in terms of the belief function $\beta_\Xi$ [6], [11], defined for any subset $S$ of $E$ by

$$\beta_\Xi(S) = P(\{\omega : \Xi(\omega) \subseteq S\}).$$

The simplest class of RFSs are the Poisson point processes. A Poisson point process is a RFS $\Upsilon$ with the property that for any $k$ disjoint subsets $S_1, ..., S_k$, the random variables $|\Upsilon \cap S_1|, ..., |\Upsilon \cap S_k|$ are independent and Poisson. Let $v_\Upsilon(S)$ denote the mean of the Poisson random variable $|\Upsilon \cap S|$, then $v_\Upsilon$ defines a (dimensionless) measure on the subsets of $E$, and is called the intensity measure of $\Upsilon$ [12]. The probability distribution $P_\Upsilon$ is given by

$$P_\Upsilon(\mathcal{T}) = e^{-v_\Upsilon(E)} \sum_{i=0}^{\infty} \frac{v_\Upsilon^i(\mathcal{T} \cap E^i)}{i!},$$

where $v_\Upsilon^i$ denotes the $i$th product measure of $v_\Upsilon$ [4]. For any subset $S$ of $E$, let $\lambda_K(S)$ denote the hyper-volume of $S$ in units of $K$. The density of $v_\Upsilon$ w.r.t. $\lambda_K$ is called the intensity function or rate of $\Upsilon$ and has units of $K^{-1}$. A Poisson point process is completely characterised by its intensity measure (or equivalently its intensity function).

A direct extension of Bayesian reasoning to multi-target systems can be achieved by interpreting the density $p_\Xi$ of a RFS $\Xi$ as the Radon-Nikodym derivative of the corresponding probability distribution $P_\Xi$ with respect to an appropriate dominating measure $\mu$, i.e

$$P_\Xi(\mathcal{T}) = \int_{\mathcal{T}} p_\Xi(X) \mu(dX).$$

The dominating measure that is often used is an unnormalised distribution of a Poisson point process with a uniform rate of $K^{-1}$ (intensity measure $\lambda = \lambda_K/K$) [4], i.e.

$$\mu(\mathcal{T}) = \sum_{i=0}^{\infty} \frac{\lambda^i(\mathcal{T} \cap E^i)}{i!}, \qquad (5)$$

For any $\mathcal{U} \subseteq \mathcal{F}(E_s)$, $\mathcal{V} \subseteq \mathcal{F}(E_o)$ let $P_{k|k}(\mathcal{U}|Z_{0:k}) \equiv P(\Xi_k \in \mathcal{U}|Z_{0:k})$ denote the posterior probability measure, $P_{k|k-1}(\mathcal{U}|X_{k-1}) \equiv P(\Xi_k \in \mathcal{U}|X_{k-1})$ and $P_k(\mathcal{V}|X_k) \equiv P(\Sigma_k \in \mathcal{V}|X_k)$ denote the conditional probability measures which describes the multi-target Markov motion and measurement respectively. Then the multi-target posterior density $p_{k|k}(\cdot|Z_{0:k})$, transition density $f_{k|k-1}(\cdot|X_{k-1})$ and likelihood $g_k(\cdot|X_k)$ used in (3-4) are the Radon-Nikodym derivatives of $P_{k|k}(\cdot|Z_{0:k})$, $P_{k|k-1}(\cdot|X_{k-1})$ and $P_k(\cdot|X_k)$ respectively.

---

[1] Where appropriate, subsets are assumed to be Borel measurable.

## 2.2. Finite Set Statistics

In this section, key concepts in finite set statistics (FISST) and its relationship with conventional point process theory are highlighted. In particular, we outline how the conditional densities $f_{k|k-1}(\cdot|\cdot)$ and $g_k(\cdot|\cdot)$ used in the recursion (3-4) can be systematically constructed from the underlying physical model of the sensors, individual target dynamics, target births and deaths using FISST.

Individual target motion in a multi-target problem is often modelled by a transition density on the state space $E_s$ while the measurement process is modelled as a likelihood on the observation space $E_o$. Consequently, the multi-target transition density and likelihood are difficult to derive since we need to compute Radon-Nikodym derivatives of probability measures on the subsets of $\mathcal{F}(E_s)$ and $\mathcal{F}(E_o)$. On the other hand, belief functions are defined directly on the subsets of $E_s$ and $E_o$. Hence, models for multi-target motion and measurement of the form

$$
\begin{aligned}
\beta_{k|k-1}(S|X_{k-1}) &\equiv P(\Xi_k \subseteq S|X_{k-1}) \\
\beta_k(T|X_k) &\equiv P(\Sigma_k \subseteq T|X_k)
\end{aligned}
$$

can be systematically constructed (see [9]). However, as belief functions are non-additive, their Radon-Nikodym derivatives (or densities) are not defined. FISST [6], [9] provides an alternative notion of density for a RFS based on the set-derivative of its belief function [6].

Let $\mathcal{B}(E)$ denote the collection of subsets of $E$. The set-derivative of a function $F : \mathcal{B}(E) \to [0, \infty)$ at a point $x \in E$ is a mapping $(dF)_x : \mathcal{B}(E) \to [0, \infty)$ defined as [6]

$$
(dF)_x(T) \equiv \lim_{\lambda_K(\Delta_x) \to 0} \frac{F(T \cup \Delta_x) - F(T)}{\lambda_K(\Delta_x)},
$$

where $\Delta_x$ denotes a neighbourhood of $x$, and $\lambda_K(\Delta_x)$ is its hyper-volume in units of $K$. This is a simplified version of the complete definition given in [6]. The set derivative at a finite set $\{x_1, ..., x_n\}$ is defined by the recursion

$$
(dF)_{\{x_1,...,x_n\}}(T) \equiv (d(dF)_{\{x_1,...,x_{n-1}\}})_{x_n}(T),
$$

where $(dF)_\emptyset \equiv F$ by convention.

For a function $f = dF_{(\cdot)}(T)$, the set-integral over a subset $S \subseteq E$ is defined as follows [6], [9]

$$
\int_S f(X)\delta X \equiv \sum_{i=0}^{\infty} \frac{1}{i!} \int_{S^i} f(\{x_1, ..., x_i\}) \lambda_K^i(dx_1...dx_i).
$$

Central to FISST is the generalised fundamental theorem of calculus

$$
f(X) = (dF)_X(\emptyset) \Leftrightarrow F(S) = \int_S f(X)\delta X.
$$

In the FISST framework, the optimal Bayes multi-target filter has the same form as (3-4) where integrals are replaced by set-integrals and densities are replaced by set-derivatives of corresponding belief functions [9].

Note that $(dF)_X(\emptyset)$ has unit of $K^{-|X|}$. Hence, the belief density evaluated at two sets $X_1$, $X_2$ with different number of elements would have different units. This lead to problems in obtaining the maximum a posteriori estimate since the units of the posterior density are not commensurable [9].

We propose a unitless set-derivative by replacing $\lambda_K$ in the definition of set-derivatives and set-integrals with the intensity measure $\lambda$ of a unit rate Poisson point process. This is equivalent to introducing a unit cancelling constant. It is easy to see that the generalised fundamental theorem of calculus still holds. Moreover, using the dominating measure given by Eq. (5), we have for any $\mathcal{U} = \cup_{i=0}^{\infty} S^i$,

$$
\int_{\mathcal{U}} f(X)\mu(dX) = \int_S f(X)\delta X. \tag{6}
$$

It follows from Eq. (6) that

$$
(d\beta_\Xi)_{(\cdot)}(\emptyset) = dP_\Xi/d\mu = p_\Xi \tag{7}
$$

i.e. the set-derivative of the belief function $\beta_\Xi$ is the density $p_\Xi$ of the corresponding probability measure $P_\Xi$ with respect to the dominating measure $\mu$ (to see this, note that for any $\mathcal{U} = \cup_{i=0}^{\infty} S^i$, $\int_S p_\Xi(X)\delta X = \int_{\mathcal{U}} p_\Xi(X)\mu(dX) = P_\Xi(\mathcal{U}) = \beta_\Xi(S)$). Consequently,

$$
\begin{aligned}
f_{k|k-1}(X_k|X_{k-1}) &= (d\beta_{k|k-1}(\cdot|X_{k-1}))_{X_k}(\emptyset), \\
g_k(Z_k|X_k) &= (d\beta_{k|k}(\cdot|X_k))_{Z_k}(\emptyset).
\end{aligned}
$$

FISST converts the construction of multi-target densities from multi-target models into computing set-derivatives of belief functions. Procedures for analytically differentiating belief functions have also been developed [6], [9] to facilitate the task for practising tracking engineers.

In general, the multi-target dynamic model (1-2) yields the following multi-target Markov transition and likelihood

$$
\begin{aligned}
&f_{k|k-1}(X_k|X_{k-1}) \\
&= \sum_{W \subseteq X_k} s_{k|k-1}(W|X_{k-1}) n_{k|k-1}(X_k - W|X_{k-1}) \quad (8)
\end{aligned}
$$

$$
g_k(Z_k|X_k) = \sum_{W \subseteq Z_k} \theta_k(W|X_k) c_k(Z_k - W|X_k). \tag{9}
$$

where $s_{k|k-1}(\cdot|X_{k-1})$ is the density of the RFS $S_k(X_{k-1})$ of surviving targets, $n_{k|k-1}(\cdot|X_{k-1})$ is the density of the RFS $N_k(X_{k-1})$ of new-born targets, $\theta_k(\cdot|X_k)$ is the density of the RFS $\Theta_k(X_k)$ of target generated observations and $c_k(\cdot|X_k)$ is the density of the RFS $C_k(X_k)$ of false alarms. Note that the difference operation used in (8-9) is the set difference. The reader is referred to [6], [9] for details on how $s_{k|k-1}(\cdot|X_{k-1})$, $n_{k|k-1}(\cdot|X_{k-1})$, $\theta_k(\cdot|X_k)$ and $c_k(\cdot|X_k)$ can be derived from the underlying physical model of the sensors, individual target dynamics, target births and deaths.

## 2.3. The PHD Filter

The probability hypothesis density (PHD) filter is a 1st order approximation of the Bayes multi-target filter. It provides a much cheaper alternative by propagating the 1st moment in stead of the full multi-target posterior [10].

A finite subset $X \in \mathcal{F}(E)$ can also be equivalently represented by a generalised function $\sum_{x \in X} \delta_x$, where $\delta_x$ denotes the Dirac delta function centred at $x$. Consequently, the random finite set $\Xi$ can also be represented by a random density $\sum_{x \in \Xi} \delta_x$. These representations are commonly used in the point process literature [1], [12].

Using the random density representation, the 1st moment or (PHD) $D_\Xi$ of a RFS $\Xi$ is defined by

$$D_\Xi(x) \equiv \mathbf{E}\left[\sum_{y \in \Xi} \delta_y(x)\right] = \int \sum_{y \in X} \delta_y(x) P_\Xi(dX). \quad (10)$$

The PHD $D_\Xi$ is a unique function (except on a set of measure zero) on the space $E$. Given a subset $S \subseteq E$, the PHD measure of $S$, i.e. $\int_S D_\Xi(x)\lambda(dx)$, gives the expected number of elements of $\Xi$ that are in $S$. The peaks of the PHD provide estimates for the elements of $\Xi$.

Let $\gamma_k$ denote PHD of the RFS $\Gamma_k$ of targets which appear spontaneously; $b_{k|k-1}(\cdot | \xi)$ denote the PHD of the RFS $B_{k|k-1}(\{\xi\})$ spawned by a target with previous state $\xi$; $e_{k|k-1}(\xi)$ denote the probability that the target still exist at time $k$ given that it has previous state $\xi$; $f_{k|k-1}(\cdot | \cdot)$ denote the transition probability density of individual targets; $g_k(\cdot | \cdot)$ denote the likelihood of individual targets; $c_k$ denote clutter probability density; $\lambda_k$ denote average number of Poisson clutter points per scan; and $p_D$ denote probability of detection. Define the PHD prediction and update operators $\Phi_{k|k-1}$, $\Psi_k$ respectively as

$$(\Phi_{k|k-1}\alpha)(x) = \int \phi_{k|k-1}(x, \xi)\alpha(\xi)\lambda(d\xi) + \gamma_k(x), \quad (11)$$

$$(\Psi_k\alpha)(x) = \left[\upsilon(x) + \sum_{z \in Z_k} \frac{\psi_{k,z}(x)}{\kappa_k(z) + \langle \psi_{k,z}, \alpha \rangle}\right]\alpha(x), \quad (12)$$

for any integrable function $\alpha$ on $E_s$, where

$$\begin{aligned}
\phi_{k|k-1}(x, \xi) &= e_{k|k-1}(\xi)f_{k|k-1}(x | \xi) + b_{k|k-1}(x | \xi), \\
\upsilon(x) &= 1 - p_D(x), \\
\psi_{k,z}(x) &= p_D(x)g_k(z | x), \\
\kappa_k(z) &= \lambda_k c_k(z).
\end{aligned}$$

Let $D_{k|k}$, denote the PHD of the multi-target posterior $p_{k|k}$. Assuming that the RFS involved are Poisson, it was shown in [10] that the PHD recursion is given by

$$D_{k|k} = (\Psi_k \circ \Phi_{k|k-1})(D_{k-1|k-1}). \quad (13)$$

Since the PHD is a function defined on the space where individual targets live, its propagation requires much less computational power than the multi-target posterior.

## 3. SEQUENTIAL MONTE CARLO IMPLEMENTATIONS

The propagation (3-4) of the multi-target posterior density recursively in time involves the evaluation of multiple set integrals and hence the computational requirement is much more intensive than single-target filtering. Sequential Monte Carlo (or particle) filtering techniques permits recursive propagation of the set of weighted random samples that approximate the full posterior [3].

### 3.1. SMC implementation of the optimal Multi-target filter

In the FISST framework, what does it mean to sample from a belief density? and how can we approximate a belief density by random samples?. The answers to these questions lie in Eq. (7) i.e. the belief density of a RFS in FISST is indeed a density of the corresponding probability distribution of the RFS (see Section 2.2). Sampling from a belief density is equivalent to sampling from the probability density and approximating a belief density is the same as approximating a probability density.

The single-target particle filter can be directly generalised to the multi-target case. In the multi-target context, however, each particle is a finite set and the particles themselves can thus be of varying dimensions.

Assume a set of weighted particles $\{w_{k-1}^{(i)}, X_{k-1}^{(i)}\}_{i=1}^N$ representing the multi-target posterior $p_{k-1|k-1}$ is available. The particle filter proceeds to approximate the multi-target posterior $p_{k|k}$ at time $k$ by a new set of weighted particles $\{w_k^{(i)}, X_k^{(i)}\}_{i=1}^N$ as follows

---

**Particle Multi-target Filter**

At time $k \geq 1$,
  **Step 1**: *Sampling Step*

- For $i = 1, ..., N$, sample $\widetilde{X}_k^{(i)} \sim q_k\left(\cdot | X_{k-1}^{(i)}, Z_k\right)$ and set

$$\widetilde{w}_k^{(i)} = \frac{g_k\left(Z_k | \widetilde{X}_k^{(i)}\right) f_{k|k-1}\left(\widetilde{X}_k^{(i)} | X_{k-1}^{(i)}\right)}{q_k\left(\widetilde{X}_k^{(i)} | X_{k-1}^{(i)}, Z_k\right)} w_{k-1}^{(i)}.$$

- Normalise weights: $\sum_{i=1}^N \widetilde{w}_k^{(i)} = 1$.

**Step 2**: *Resampling Step*

- Resample $\left\{\widetilde{w}_k^{(i)}, \widetilde{X}_k^{(i)}\right\}_{i=1}^N$ to get $\left\{w_k^{(i)}, X_k^{(i)}\right\}_{i=1}^N$.

---

The importance sampling density $q_k(\cdot | X_{k-1}, Z_k)$ is a multi-target density and $\widetilde{X}_k^{(i)} \sim q_k(\cdot | X_{k-1}^{(i)}, Z_k)$ is a sample from a RFS or point process. Sampling from a point

process is well studied see for example [1], [12] and references therein.

After the resampling step, an optional MCMC step can also be applied to increase particle diversity [5]. Since the particles reside in spaces of different dimensions, a reversible jump MCMC step [7] is required. Under standard assumptions, convergence results for particle filters also apply to the multi-target case [2].

From the multi-target posterior $p_{k|k}$, an estimate of the target set at time $k$ can be given by the expected a posteriori estimator [6], [9]. This estimator is the 1st moment (or PHD) $D_{k|k}$ (see Section 2.3 Eq. (10)). Given a particle approximation $\{w_k^{(i)}, X_k^{(i)}\}_{i=1}^{N}$ of $p_{k|k}$, the particle approximation of $D_{k|k}$ is then

$$
\begin{aligned}
D_{k|k}(x) &= \int \sum_{y \in X} \delta_y(x) P_{k|k}(dX|Z_{0:k}) \\
&\approx \sum_{i=1}^{N} \sum_{y \in X_k^{(i)}} w_k^{(i)} \delta_y(x)
\end{aligned}
$$

The main practical problem with the multi-target particle filter is the need to perform importance sampling in very high dimensional spaces if many targets are present. Moreover, it can be difficult to find an efficient importance density and the choice of a naive importance density like $q_k(\cdot \,|\, X_{k-1}^{(i)}, Z_k) = f_{k|k-1}(\cdot \,|\, X_{k-1}^{(i)})$ will typically lead to an algorithm whose efficiency decreases exponentially with the number of targets for a fixed number of particles.

### 3.2. SMC implementation of the PHD filter

The PHD filter is a cheaper alternative. However, direct application of standard SMC methods to propagate the multi-target PHD would fail because firstly, the PHD is not a probability density function; and secondly, it is not a standard Bayes recursion. In this section we summarise the SMC implementation of the PHD filter proposed in [13].

For any $k \geq 0$, let $\alpha_k = \{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{L_k}$ denote a particle approximation of $D_{k|k}$. Using the PHD recursion, a particle approximation of the PHD at time step $k > 0$ can be obtained from a particle approximation at the previous time step by the following procedure (see [13] for the derivation)

---

**SMC PHD filter**

At time $k \geq 1$,
    **Step 1**: _Prediction_

- Sample

$$
\widetilde{x}_k^{(i)} \sim \begin{cases} q_k\left(\cdot \,|\, x_{k-1}^{(i)}, Z_k\right), \ i = 1, ..., L_{k-1} \\ p_k\left(\cdot \,|\, Z_k\right), \ i = L_{k-1}+1, ..., L_{k-1}+J_k \end{cases}
$$

---

and compute the predicted weights

$$
\widetilde{w}_{k|k-1}^{(i)} = \begin{cases} \dfrac{\phi_k\left(\widetilde{x}_k^{(i)}, x_{k-1}^{(i)}\right) w_{k-1}^{(i)}}{q_k\left(\widetilde{x}_k^{(i)} \,|\, x_{k-1}^{(i)}, Z_k\right)}, \ i = 1, ..., L_{k-1} \\ \dfrac{\gamma_k\left(\widetilde{x}_k^{(i)}\right)}{J_k p_k\left(\widetilde{x}_k^{(i)} \,|\, Z_k\right)}, \ i = L_{k-1}+1, ..., L_{k-1}+J_k \end{cases}
$$

**Step 2:** _Update_

- For each $z \in Z_k$, compute

$$
C_k(z) = \sum_{j=1}^{L_{k-1}+J_k} \psi_{k,z}(\widetilde{x}_k^{(j)}) \widetilde{w}_{k|k-1}^{(j)}.
$$

- For $i = 1, ..., L_{k-1}+J_k$, update weights

$$
\widetilde{w}_k^{(i)} = \left[ \upsilon(\widetilde{x}_k^{(i)}) + \sum_{z \in Z_k} \frac{\psi_{k,z}(\widetilde{x}_k^{(i)})}{\kappa_k(z) + C_k(z)} \right] \widetilde{w}_{k|k-1}^{(i)}.
$$

**Step 3:** _Resampling_

- Compute the total mass $\widehat{N}_{k|k} = \sum_{j=1}^{L_{k-1}+J_k} \widetilde{w}_k^{(j)}$,

- Resample $\left\{ \dfrac{\widetilde{w}_k^{(i)}}{\widehat{N}_{k|k}}, \widetilde{x}_k^{(i)} \right\}_{i=1}^{L_{k-1}+J_k}$ to get $\left\{ \dfrac{w_k^{(i)}}{\widehat{N}_{k|k}}, x_k^{(i)} \right\}_{i=1}^{L_k}$,

- Rescale the weights by $\widehat{N}_{k|k}$ to get $\left\{ w_k^{(i)}, x_k^{(i)} \right\}_{i=1}^{L_k}$.

---

Note that the computational complexity of this algorithm is independent of the (time-varying) number of targets if we fix the number of particles $L_k = L$ for all $k$. However, this can result in instances where the number of particles are not sufficient to resolve a large number of targets while at other times we may have an excess of particles for a small number of targets. On the other hand, if $L_k = L_{k-1} + J_k$, then $L_k$ increases over time even if the number of targets does not. This is very inefficient, since computational resource is wasted in exploring regions of the state space where there are no targets. It would be computationally more efficient to adaptively allocate approximately say $\rho$ particles per target at each time step.

Since the expected number of targets $\int D_{k|k}(\xi|Z_{0:k})d\xi$ can be estimated by $\widehat{N}_{k|k} = \sum_{j=1}^{L_{k-1}+J_k} w_k^{(j)}$, it is natural to have the number of particles $L_k \cong \rho \widehat{N}_{k|k}$. This can be achieved by resampling $L_k \cong \rho \widehat{N}_{k|k}$ particles from $\{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{L_{k-1}+J_k}$ and redistributing the total mass $\widehat{N}_{k|k}$ among the $L_k$ resampled particles.

In the resampling step of the SMC PHD filter, the new weights $\{w_k^{(i)}\}_{i=1}^{L_k}$ are not normalised to 1 (as in conventional particle filter) but sum to $\widehat{N}_{k|k}$ instead. Similarly to the standard case, each particle $\widetilde{x}_k^{(i)}$ is copied $N_k^{(i)}$ times under the constraint $\sum_{i=1}^{L_{k-1}+J_k} N_k^{(i)} = L_k$ to obtain $\{x_k^{(i)}\}_{i=1}^{L_k}$. The (random) resampling mechanism is chosen such that $E[N_k^{(i)}] = L_k a_k^{(i)}$ where $a_k^{(i)} > 0$, $\sum_{i=1}^{L_{k-1}+J_k} a_k^{(i)} = 1$ is a sequence of weights set by the user. However, the new weights are set to $w_k^{(i)} \propto \widetilde{w}_k^{(i)}/a_k^{(i)}$ with $\sum_{i=1}^{L_k} w_k^{(i)} = \widehat{N}_{k|k}$ instead of $\sum_{i=1}^{L_k} w_k^{(i)} = 1$. Typically, $a_k^{(i)} = \widetilde{w}_k^{(i)}/\widehat{N}_{k|k}$.

## 4. SIMULATIONS

For visualisation purposes, consider a one-dimensional scenario with an unknown and time varying number of targets observed in clutter. The states of the targets consist of 1-D position and 1-D velocity, while only position measurements are available. Without loss of generality, we consider targets with linear Gaussian dynamics. We assume a Poisson model for spontaneous target birth with intensity $0.2N(\cdot\,|\,0,1)$. Each existing a target has a (state independent) probability of survival $e = 0.8$. The clutter process is Poisson with uniform intensity over the region $[-100; 100]$ and has an average rate of 10. Figure 1 shows the tracks with clutter on the position measurements and Figure 2 plots the PHD of position against time. Observe from Figure 2 that the PHD filter shows surprisingly good performance, even very short tracks are picked up among clutter.
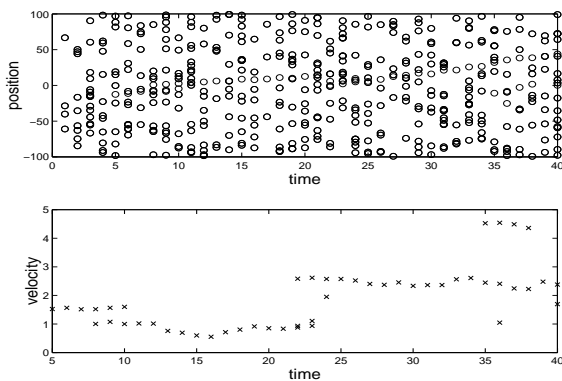


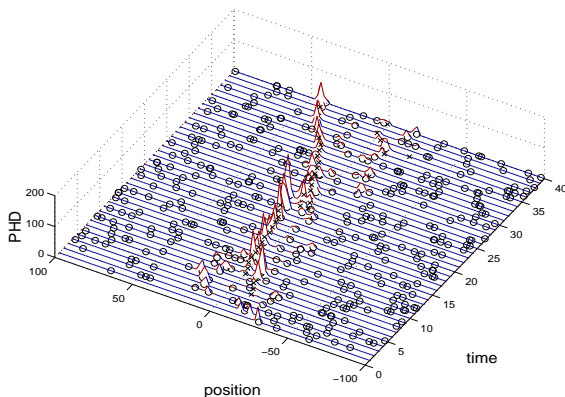**Fig. 1**. birth rate 0.2, death rate 0.2, clutter rate 10



**Fig. 2**. PHD of position

## 5. CONCLUSION

In this paper, we have highlighted the relationship between Radon-Nikodym derivative and set derivative of random fi-

nite sets. This allows the derivation of a generic particle filter for estimating an evolving random finite set. In addition, a sequential Monte Carlo implementation of the probability hypothesis density filter for multi-target tracking is also given. With a computational complexity that is independent of the (time-varying) number of targets, this offers a very practical alternative to the particle random set filter.

## 6. REFERENCES

[1] D. Daley and D. Vere-Jones, *An introduction to the theory of point processes*, Springer-Verlag, 1988.

[2] P. Del Moral and L. Miclo "Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae with Applications to Non-Linear Filtering", *Seminaire de Probabilites XXXIV*, J. Azima et al. (Eds.), Lecture Notes in Mathematics, Springer-Verlag Berlin, Vol. 1729, 1-145 (2000).

[3] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, 2001.

[4] C. Geyer, "Likelihood inference for spatial point processes", in *Stochastic Geometry likelihood and computation*, Barndorff-Nielsen et. al (eds.), pp. 79-140, 1999.

[5] W. Gilks and C. Berzuini, "Following a moving target-MonteCarlo inference for dynamic Bayesian models", Journal of the Royal Statistical Society, Series B, 63, 2001.

[6] I. Goodman, R. Mahler and H. Nguyen, *Mathematics of Data Fusion*, Kluwer Academic Publishers, 1997.

[7] P. J. Green, "Reversible jump MCMC computation and Bayesian model determination", *Biometrika*, 82: 711-732, 1995.

[8] R. Mahler. "Global integrated data fusion" , *Proc. 7th Nat. Symp. on Sensor Fusion*, Vol. 1, (Unclassified) Sandia National Laboratories, Albuquerque, ERIM Ann Arbor MI, pp. 187-199, 1994.

[9] R. Mahler, *An Introduction to Multisource-Multitarget Statistics and Applications*, Lockheed Martin Technical Monograph. March 15, 2000.

[10] R. Mahler, "Approximate multisensor-multitarget joint detection, tracking and identification using a 1st order multitarget moment statistic," *IEEE Trans. AES*, to appear.

[11] G. Matheron, *Random sets and integral geometry*, J. Wiley, 1975,

[12] D. Stoyan, D. Kendall, and J. Mecke, *Stochastic Geometry and its applications*, John Wiley & Sons, 1995.

[13] B. Vo, S. Singh and A. Doucet, "Sequential Monte Carlo Implementation of the PHD Filter for Multi-target Tracking", *Proc. FUSION'2003*, Cairns, Australia.