Sequential Monte Carlo Implementation of the PHD Filter for Multi-target Tracking

Ba-Ngu Vo, Sumeetpal Singh
Department of Electrical & Electronic Engineering,
The University of Melbourne,
Victoria, 3010, Australia
Email: {bv,ssss}@ee.mu.oz.au

Arnaud Doucet
Department of Engineering,
Cambridge University,
CB2 1PZ Cambridge, UK
Email: ad2@eng.cam.ac.uk

Abstract – *Random finite sets are natural representations of multi-target states and observations that allow multi-sensor multi-target tracking to fit in the unifying random set framework for Data Fusion. Although a rigorous foundation has been developed in the form of Finite Set Statistics, optimal Bayesian multi-target filtering is not yet practical. Sequential Monte Carlo (SMC) approximations of the optimal filter are computationally expensive. A practical alternative to the optimal filter is the Probability Hypothesis Density (PHD) filter, which propagates the PHD or first moment instead of the full multi-target posterior. The propagation of the PHD involves multiple integrals which do not admit closed form. We propose to approximate the PHD by a set of weighted random samples which are propagated over time using a generalised SMC method. The resulting algorithm is very attractive as it is general enough to handle non-linear non-Gaussian dynamics and the 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

Multi-sensor multi-target tracking is a class of dynamic state estimation problems in which the entity of interest is a set that is random in the number of elements as well as the values of individual elements [2]. Finite random sets are therefore natural and intuitive representations of multi-target states and multi-target measurements. The modelling of multi-target dynamic using random sets naturally leads to tracking algorithms which incorporate track initiation, a procedure that has mostly been performed separately in traditional tracking algorithms. More importantly, random sets provide a rigorous unified framework for the seemingly unconnected sub-disciplines of data fusion [7], [11].

Although stochastic geometrical models, including deformable templates and random finite sets have long been used by statisticians to develop techniques for object recognition in static images [1], this representation has been largely overlooked in the data fusion and tracking literature until recently [10], [11]. The first systematic treatment of multi-sensor multi-target tracking, as part of a unified framework for data fusion using random set theory was the work of Mahler [10], [11]. By reconceptualising all sensors as a single meta sensor, the target set as a single meta target with multi-target state, and the observations collected by the sensor suite as a single set of measurements of the meta sensor, the multi-sensor multi-target tracking problem can be rigorously cast in a Bayesian framework.

Analogous to single-target tracking, Bayesian multi-target tracking (in the random finite set framework) propagates the multi-target posterior density recursively in time. This involves the evaluation of multiple (set) integrals and the computational intractability is more severe than its single-target counterpart. Moreover, naive Sequential Monte Carlo (SMC) methods have an efficiency decreasing exponentially with the number of targets.

The Probability Hypothesis Density (PHD) filter proposed by Mahler [12],[13] is a tractable alternative to the optimal multi-target filter. It is a recursion propagating the PHD or first moment of the multi-target posterior. Under the assumption that the predicted multi-target density is Poisson, this recursion is exact and completely characterises the statistics of the dynamic Poisson point process of interest. We emphasise here that the PHD is a function defined on the space where individual targets live, but it is not a probability density. It is positive and integrable, but does not necessarily integrate to unity. Unfortunately, the PHD propagation equations involve multiple integrals that have no computationally tractable closed form expressions even for the simple case where individual targets follow a linear Gaussian dynamic.

In this article, we propose to approximate the PHD using a SMC method; i.e. the PHD is represented by a large set of weighted random samples (or particles) which are propagated over time using importance sampling and re-sampling strategies. SMC methods are extremely powerful tools which have had an impact on optimal (Bayesian) filtering [4], [5]. However, a direct application of standard SMC methods to propagate the multi-target PHD would fail as the PHD is not a probability density function and the PHD recursion is not a standard Bayes recursion. In this paper, a probabilistic interpretation of the PHD recursion is given which allows the derivation of an efficient SMC implementation of the PHD filter. The proposed algorithm is general enough to handle non-linear non-Gaussian dynamics. The salient feature of this algorithm is that its computational complexity is independent of the time-varying number of targets, though the number of particles can be adaptively allocated in such a way that a constant ratio between the number of particles and the expected number of targets is maintained. Numerical studies shows surprisingly good performance, even very short tracks are picked up among clutter.

The rest of the paper is organised as follows. Section 2 briefly reviews the basics of random finite set, optimal multi-target tracking and the PHD recursion. Section 3 describes a particle implementation of the PHD recursion. Simulation results are presented in Section 4. Finally, some potential extensions are discussed in Section 5.

2 Random Finite Set Model of Multi-target Dynamics

In a single-target system, the state and measurement at time k are two vectors of possibly different dimensions, while in a multi-target system the state and measurement are two collections of individual targets and measurements. The state and measurement of a single-target system evolve in time with their dimensions unchanged. However, this is not the case with a multi-target problem. As the multi-target state and measurement evolve in time, the number of individual targets and measurements may change, i.e. the dimensions of the multi-target state and multi-target measurement also evolve in time.

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 located at $x_{k,1}, \ldots, x_{k,M(k)} \in E_s$ then, $X_k = \{x_{k,1}, \ldots, x_{k,M(k)}\} \subseteq E_s$. Similarly, if $N(k)$ observations $z_{k,1}, \ldots, z_{k,N(k)} \in E_o$ 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) Ξ_k and Σ_k on the state and observation spaces E_s and E_o respectively. A formal definition of a random finite set is given in Section 2.1.

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

$$\Xi_k = S_k(X_{k-1}) \cup B_k(X_{k-1}) \cup \Gamma_k$$

where $S_k(X_{k-1})$ denotes the RFS of targets that have survived at time k, $B_k(X_{k-1})$ denotes the RFS of targets spawned from X_{k-1} and Γ_k denotes the RFS of targets that appear spontaneously at time k. The statistical behaviour of the RFS Ξ_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 the multi-target state at time k, the multi-target measurement can be modelled by the RFS

$$\Sigma_k = E_k(X_k) \cup C_k(X_k)$$

where $E_k(X_k)$ denotes the RFS of measurements generated by X_k, and $C_k(X_k)$ denotes the RFS of clutter or false alarms. The statistical behaviour of the RFS Σ_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 observations.

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

$$\begin{aligned} & p_{k|k-1}(X_k|Z_{0:k-1}) \\ & \quad = \int f_{k|k-1}(X_k|X) p_{k-1|k-1}(X|Z_{0:k-1}) \mu_s(dX) \quad (1) \\ & p_{k|k}(X_k|Z_{0:k}) \\ & \quad = \frac{g_k(Z_k|X_k) p_{k|k-1}(X_k|Z_{0:k-1})}{\int g_k(Z_k|X) p_{k|k-1}(X|Z_{0:k-1}) \mu_s(dX)}. \quad (2) \end{aligned}$$

where μ_s is a dominating measure to be discussed later. The main difference between the recursion (1-2) and standard clutter-free single-target filtering is that X_k and Z_k can change dimension as k changes. In most cases, the above recursion cannot be done analytically.

In Section 2.3, we outline a sequential Monte Carlo approach to implement this recursion. Section 2.1 presents a formal definition of RFS and formalises the notion of density. In Section 2.2, we outline how $f_{k|k-1}$ and g_k may be constructed from the underlying physical model of the sensors, individual target dynamics as well as target births and deaths. Sections 2.1 and 2.2 can be omitted without loss of continuity provided the notions of RFS and its density are accepted on faith.

2.1 Random Finite Sets

For completeness, this section outlines some background materials on random finite sets (RFS) or simple point processes. Background materials on RFS are abundant in the spatial statistics literature; see for example [3], [15]. However, works with an inclination to multi-target tracking are quite new; the major body of work appears to be that of Mahler [7], [10]. The monograph [11] is an excellent introduction accessible to a wide range of readers.

Given a compact subset E of $\mathbf{R}^n$, let $\mathcal{F}(E)$ denote the collection of finite subsets of E. A random finite set Ξ on E is defined as a measurable mapping from a probability space $(\Omega, \mathcal{A}, P)$ to the collection $(\mathcal{F}(E), \mathcal{B}(\mathcal{F}))$ of finite subsets of E

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

The probability measure P on the sample space Ω induces a probability law for Ξ, which can be specified in terms of probability distribution, void probabilities or belief function. The probability distribution P_Ξ on the (Borel) subsets $\mathcal{B}(\mathcal{F})$ of $\mathcal{F}(E)$ is the most natural description of the probability law for Ξ

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

However, from random set theory [7], [14], the probability law for Ξ can also be given in terms of the belief function β_Ξ on the bounded Borel subsets of E

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

An equivalent description that is closely related to the belief function is the void probability v_Ξ [3], [15], on the bounded Borel subsets of E

$$v_\Xi(S) = P(\{\omega : |\Xi(\omega) \cap S| = 0\}) = \beta_\Xi(S^c).$$

Note that the above discussion is still valid when E is replaced by any locally compact Haussdorff separable space. We will use E to denote either the state space E_s or observation space E_o.

A direct extension of Bayesian reasoning to multi-target systems can be achieved by interpreting the probability density p_Ξ of a RFS Ξ as the Radon-Nikodym derivative of the corresponding probability distribution P_Ξ with respect to an appropriate dominating measure μ, i.e $P_\Xi(\mathcal{T}) = \int_\mathcal{T} p_\Xi(X)\mu(dX)$. In spatial statistics, the dominating measure that is often used is given by [6]

$$\mu(\mathcal{T}) = \sum_{i=0}^{\infty} \lambda^i(\mathcal{T} \cap E^i)/i!, \quad (3)$$

where λ^i denotes the Lebesgue measure on E^i.

For any Borel subsets $\mathcal{U} \subseteq \mathcal{F}(E_s)$, $\mathcal{V} \subseteq \mathcal{F}(E_o)$ 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 $p_{k|k}(X_k|Z_{0:k})$, transition density $f_{k|k-1}(X_k|X_{k-1})$ and likelihood $g_k(Z_k|X_k)$ used in (1-2) 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 i.e.

$$\begin{aligned}
P_{k|k}(\mathcal{U}|Z_{0:k}) &= \int_\mathcal{U} p_{k|k}(X_k|Z_{0:k})\mu_s(dX_k), \\
P_{k|k-1}(\mathcal{U}|X_{k-1}) &= \int_\mathcal{U} f_{k|k-1}(X_k|X_{k-1})\mu_s(dX_k), \\
P_k(\mathcal{V}|X_k) &= \int_\mathcal{V} g_k(Z_k|X_k)\mu_o(dZ_k).
\end{aligned}$$

where μ_s and μ_o are dominating measures on the Borel subsets of $\mathcal{F}(E_s)$ and $\mathcal{F}(E_o)$ respectively.

2.2 Finite Set Statistics

In this section, key concepts in finite set statistics and its relationship with conventional spatial statistics are highlighted.

Individual target motion in a multi-target problem is often modelled by a transition density on E_s while the measurement process is modelled as a likelihood on E_o. Consequently, multi-target transition density and likelihood as Radon-Nikodym derivatives (of measures on the Borel subsets of $\mathcal{F}(E_s)$ and $\mathcal{F}(E_o)$) are difficult to construct. Finite set statistics [7], [11] provides an alternative notion of density for a RFS based on belief function. Since belief functions are defined on subsets of E_s, E_o, models for multi-target motion and measurement of the form $\beta_{k|k-1}(S|X_{k-1}) \equiv P(\Xi_k \subseteq S|X_{k-1})$, $\beta_k(T|X_k) \equiv P(\Sigma_k \subseteq T|X_k)$ can be easily constructed. However, belief functions are non-additive, hence their Radon-Nikodym derivatives are not defined. This problem can be addressed by the introduction of set integrals and set derivatives [7].

For a given function $f : \mathcal{F}(E) \to [0, \infty)$, the set integral over a bounded Borel subset $S \subseteq E$ is defined as

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

Note that using the dominating measure given by Eq. (3), we have for any $\mathcal{U} = \cup_{i=0}^{\infty} S^i$,

$$\int_{\mathcal{U}} f(X)\mu(dX) = \int_S f(X)\delta X. \quad (4)$$

For a set function $F : \mathcal{B}(E) \to [0, \infty)$ its density $f : \mathcal{F}(E) \to [0, \infty)$ is defined by

$$F(S) = \int_S f(X)\delta X.$$

The set derivative of a function $F : \mathcal{B}(E) \to [0, \infty)$ at a point x is a mapping $(dF)_x : \mathcal{B}(E) \to [0, \infty)$ defined as

$$(dF)_x(S) \equiv \lim_{\lambda(\Delta_x)\to 0} \frac{F(S \cup \Delta_x) - F(S)}{\lambda(\Delta_x)}.$$

where Δ_x denotes a neighbourhood of x. Note that this is a simplified definition, see [7] for a proper definition. Furthermore, the set derivative at a finite set $X = \{x_1, ..., x_n\}$ is defined by the recursion

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

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

Central to finite set statistics is the generalised fundamental theorem of calculus

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

which allows the density of a non-additive set function to be determined constructively. It is clear from Eq. (4) that $(d\beta_\Xi)_{(\cdot)}(\emptyset) = dP_\Xi/d\mu$, i.e. the density of the belief function β_Ξ is, in fact, the density of the corresponding probability measure P_Ξ with respect to the dominating measure μ. Consequently, the multi-target transition density $f_{k|k-1}(X_k|X_{k-1})$ and likelihood $g_k(Z_k|X_k)$ can be determined explicitly by

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

Finite set statistics 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 [7], [11] to facilitate the task for practising tracking engineers.

2.3 Multi-target Particle Filter

The propagation 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. Particle filtering techniques permits recursive propagation of the full posterior [4], [5]. The single-target particle filter can be directly generalised to the multi-target case. Assume at time $k-1$, a set of weighted particles $\{w_{k-1}^{(i)}, X_{k-1}^{(i)}\}_{i=1}^N$ is available, the particle filter proceeds as follows at time k

---

Multi-target Particle filter

At time $k \geq 1$,

Step 1: Sampling Step

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

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

- 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 $\left\{w_k^{(i)}, X_k^{(i)}\right\}_{i=1}^N$.

In this context, 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 point process.

There are many possible ways to perform the resampling step. Most methods consists of copying each particle $\widetilde{X}_k^{(i)}$ $N_k^{(i)}$ times under the constraint $\sum_{i=1}^N N_k^{(i)} = N$ to obtain $\{X_k^{(i)}\}_{i=1}^N$. The (random) resampling mechanism is chosen such that $E[N_k^{(i)}] = N\alpha_k^{(i)}$ where $\alpha_k^{(i)} > 0$, $\sum_{i=1}^N \alpha_k^{(i)} = 1$ is a sequence of weights set by the user. This resampling step could be achieved using multinomial resampling but the efficient stratified resampling algorithm described in [9] has better statistical properties. The new weights are set to $w_k^{(i)} \propto \widetilde{w}_k^{(i)}/\alpha_k^{(i)}$, $\sum_{i=1}^N w_k^{(i)} = 1$. Typically, $\alpha_k^{(i)} = \widetilde{w}_k^{(i)}$ but alternatively we can select $\alpha_k^{(i)} \propto (\widetilde{w}_k^{(i)})^\nu$ where $\nu \in (0, 1)$.

In this algorithm, each particle corresponds to a sample from a point process and the particles can thus be of varying dimensions. The main practical problem with this approach is the need to perform importance sampling in very high dimensional spaces if many targets are present. Moreover, it can be difficult to come up with an efficient importance distribution and the choice of a naive importance distribution 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 would decrease exponentially with the number of targets for a fixed number of particles.

2.4 The PHD Filter

The Probability Hypothesis Density (PHD) of a RFS is the analogue of the expectation of a random vector. The expectation of a random set, however, has no meaning since there is no notion of addition for sets. Nevertheless, an indirect construction can be used by representing random sets as random counting measures or random density functions.

A finite subset $X \in \mathcal{F}(E)$ can also be equivalently represented by the counting measure N_X (on the Borel subsets of E) defined by $N_X(S) = \sum_{x\in X} 1_S(x) = |X \cap S|$, where the notation $|A|$ denotes the number of elements in A. Alternatively, the density δ_X of N_X can also be used to represent the finite set X. Indeed, $\delta_X = \sum_{x\in X} \delta_x$, where δ_x denotes the Dirac delta function centred at x. Consequently, the random finite set Ξ can also be represented by a random counting measure N_Ξ defined by $N_\Xi(S) = |\Xi \cap S|$ or its random density $\delta_\Xi = \sum_{x\in\Xi} \delta_x$. These representations are commonly used in the point process literature [3], [15].

Using the random density representation, the first order moment (or PHD) D_Ξ of a RFS Ξ is defined by

$$D_\Xi(x) \equiv \mathbf{E}[\delta_\Xi(x)] = \int \delta_X(x) P_\Xi(dX).$$

It can be shown that [12],[13]

$$D_\Xi(x) = \int f_\Xi(\{x\} \cup W)\delta W = \quad (d\beta_\Xi)_x(E).$$

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

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

$$(\Phi_{k|k-1}\alpha)(x) = \int \phi_{k|k-1}(x,\xi)\alpha(\xi)\lambda(d\xi) + \gamma_k(x), \quad (5)$$

$$(\Psi_k\alpha)(x) = \left[\upsilon(x) + \sum_{z\in Z_k} \frac{\psi_{k,z}(x)}{\kappa_k(z) + \langle\psi_{k,z},\alpha\rangle}\right]\alpha(x), \quad (6)$$

for any integrable function α 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 [12],[13] that the PHD recursion is given by

$$D_{k|k} = \left(\Psi_k \circ \Phi_{k|k-1}\right)\left(D_{k-1|k-1}\right). \quad (7)$$

Observe that the prediction operator $\Phi_{k|k-1}$ is affine while the update operator Ψ_k is highly non-linear. The update operator appears to be linear at first glance since it only scales the argument. However, the scaling factor is a highly non-linear function of the argument.

3 Sequential Monte Carlo Implementation of the PHD Filter

In this section, we detail a novel Sequential Monte Carlo (SMC) method to implement the recursion described by (5)-(6). A particle interpretation of the prediction and update operators is presented first followed by the full algorithm.

3.1 The prediction operator

Suppose that at time step $k-1$, we have a function α_{k-1} characterised (exactly) by the set of particles and weights $\{w_{k-1}^{(i)}, x_{k-1}^{(i)}\}_{i=1}^{L_{k-1}}$ i.e.

$$\alpha_{k-1}(x) = \sum_{i=1}^{L_{k-1}} w_{k-1}^{(i)}\delta_{x_{k-1}^{(i)}}(x)$$

(the notation $\{w_{k-1}^{(i)}, x_{k-1}^{(i)}\}_{i=1}^{L_{k-1}}$ is used interchangeably with $\sum_{i=1}^{L_{k-1}} w_{k-1}^{(i)}\delta_{x_{k-1}^{(i)}}$). Then

$$\begin{aligned}
(\Phi_{k|k-1}\alpha_{k-1})(x_k) &= \int \phi_{k|k-1}(x_k,\xi)\alpha_{k-1}(\xi)\lambda(d\xi)\\
&+ \gamma_k(x_k)\\
&= \sum_{i=1}^{L_{k-1}} w_{k-1}^{(i)}\phi_k(x_k, x_{k-1}^{(i)}) + \gamma_k(x_k)
\end{aligned}$$

To obtain a particle approximation of $\Phi_{k|k-1}\alpha_{k-1}$, we apply importance sampling to each of its terms. Let $\{x_k^{(i)}\}_{i=1}^{L_{k-1}}$ be L_{k-1} samples from the proposal density

$q_k(\cdot|\, x_{k-1}^{(i)}, Z_k)$ and $\{x_k^{(i)}\}_{i=L_{k-1}+1}^{L_{k-1}+J_k}$ be J_k i.i.d. samples from another proposal density $p_k(\cdot|\, Z_k)$ i.e.

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

We can now define the approximate prediction operator $\widehat{\Phi}_{k|k-1}$ that maps the particle representation $\{w_{k-1}^{(i)}, x_{k-1}^{(i)}\}_{i=1}^{L_{k-1}}$ to another particle representation $\{w_{k|k-1}^{(i)}, x_k^{(i)}\}_{i=1}^{L_{k-1}+J_k}$ as follows

$$(\widehat{\Phi}_{k|k-1}\alpha_{k-1})(x_k) \equiv \sum_{i=1}^{L_{k-1}+J_k} w_{k|k-1}^{(i)} \delta_{x_k^{(i)}}(x_k)$$

where

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

Note that we started with α_{k-1} having L_{k-1} particles, which are then predicted forward by the kernel $\phi_{k|k-1}$ to another set of L_{k-1} particles. Additionally, we also have J_k new particles arising from the birth process. The number of new particles J_k can be a function of k to accommodate the varying number of new born targets at each time step. Assuming that the total mass of γ_k has closed form, then typically J_k is chosen to be proportional to this mass, i.e. $J_k = \rho \int \gamma_k(x)dx$, so that on average we have ρ particles per new born target.

3.2 The update operator

For the update step of the recursion, assume that we have from the prediction a function $\alpha_{k|k-1}$ characterised by $\{w_{k|k-1}^{(i)}, x_k^{(i)}\}_{i=1}^{L_{k-1}+J_k}$. Applying the update operator gives

$$(\Psi_k \alpha_{k|k-1})(x) = \sum_{i=1}^{L_{k-1}+J_k} w_k^{(i)} \delta_{x_k^{(i)}}(x),$$

where

$$w_k^{(i)} = \left[\upsilon(x^{(i)}) + \sum_{z\in Z_k} \frac{\psi_{k,z}(x_k^{(i)})}{\kappa_k(z) + C_k(z)} \right] w_{k|k-1}^{(i)}, \quad (8)$$

$$C_k(z) = \sum_{j=1}^{L_{k-1}+J_k} \psi_{k,z}(x_k^{(j)}) w_{k|k-1}^{(j)}. \quad (9)$$

The update operator maps the function with particle representation $\{w_{k|k-1}^{(i)}, x_k^{(i)}\}_{i=1}^{L_{k-1}+J_k}$ into one with particle representation $\{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{L_{k-1}+J_k}$ by modifying the weights of these particles according to Eq. (8).

3.3 Particle propagation

For any $k \geq 0$, let $\alpha_k = \{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{L_k}$ denote a particle approximation of $D_{k|k}$. The algorithm is designed such that the concentration of particles in a given region of the state space, say A, represents the expected number of targets in A.

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

$$\alpha_k = (\Psi_k \circ \widehat{\Phi}_{k|k-1})\alpha_{k-1}. \quad (10)$$

Note that since α_k has $L_k = L_{k-1} + J_k$ particles, the number L_k of particles 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. On the other hand if L_k is fixed then the ratio of particles to targets would fluctuate as the number of targets changes. In other words, at times we may have an insufficient number of particles to resolve the targets (up to the PHD limitations) while at other times we may have an excess of particles for a small number of targets or no target at all. It would be computationally more efficient to adaptively allocate approximately say ρ particles per target at each time epoch.

Since the expected number of targets $N_{k|k}$ (given by the total mass $\int D_{k|k}(\xi|Z_{0:k})d\xi$) 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 $L_k \cong \rho \widehat{N}_{k|k}$. Furthermore, we also want to eliminate particles with low weights and multiply particles with high weights to focus on the important zones of the space. 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.

3.4 Algorithm

Based on the elements presented above, it is possible to propose the following generic particle filtering algorithm for the operator recursion.

Particle PHD filter

At time $k \geq 1$,

Step 1: <u>*Prediction step*</u>

- For $i = 1, ..., L_{k-1}$, sample $\widetilde{x}_k^{(i)} \sim q_k\left(\cdot|\, x_{k-1}^{(i)}, Z_k\right)$ and compute the predicted weights

$$\widetilde{w}_{k|k-1}^{(i)} = \frac{\phi_k(\widetilde{x}_k^{(i)}, x_{k-1}^{(i)})}{q_k\left(\widetilde{x}_k^{(i)} \middle| x_{k-1}^{(i)}, Z_k\right)} w_{k-1}^{(i)}.$$

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

$$\widetilde{x}_k^{(i)} \sim p_k\left(\cdot \mid Z_k\right)$$

and compute the weights of new born particles

$$\widetilde{w}_{k|k-1}^{(i)} = \frac{1}{J_k} \frac{\gamma_k\left(\widetilde{x}_k^{(i)}\right)}{p_k\left(\widetilde{x}_k^{(i)} \middle| Z_k\right)}.$$

Step 2: *Update step*

- 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 step*

- Compute the total mass $\widehat{N}_{k|k} = \sum_{j=1}^{L_{k-1}+J_k} \widetilde{w}_k^{(j)}$
- Resample $\left\{\widetilde{w}_k^{(i)}/\widehat{N}_{k|k}, \widetilde{x}_k^{(i)}\right\}_{i=1}^{L_{k-1}+J_k}$ to get $\left\{w_k^{(i)}/\widehat{N}_{k|k}, x_k^{(i)}\right\}_{i=1}^{L_k}$.

Care must be taken when implementing the resampling step for the particle PHD filter. In this case, the new weights $\{w_k^{(i)}\}_{i=1}^{L_k}$ are not normalised to 1 but sum to $\widehat{N}_{k|k}$. 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 $E[N_k^{(i)}] = L_k \alpha_k^{(i)}$ where $\alpha_k^{(i)} > 0$, $\sum_{i=1}^{L_{k-1}+J_k} \alpha_k^{(i)} = 1$ is a sequence of weights set by the user. This is achieved using stratified resampling [9]. However, the new weights are set to $w_k^{(i)} \propto \widetilde{w}_k^{(i)}/\alpha_k^{(i)}$ 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, $\alpha_k^{(i)} = \widetilde{w}_k^{(i)}/\widehat{N}_{k|k}$ but alternatively we can select $\alpha_k^{(i)} \propto (\widetilde{w}_k^{(i)})^{\nu}$ where $\nu \in (0, 1)$.

This filter reduces to the standard particle filter in the case where there is only one target with no birth, no death and no clutter.

In the standard particle filtering context, choosing the importance distribution so as to minimise the (conditional) variance of the weights, is well known. In the PHD context, this becomes much more difficult and is the subject of further study.

4 Simulations

For visualisation purposes, a one-dimensional scenario is considered. The targets move along the line segment $[-100; 100]$. The states of the targets consist of position and velocity, while only position measurements are obtained. Targets can appear or disappear in the scene at any time. We assume a Poisson model for spontaneous target birth. Without loss of generality, we consider targets with linear Gaussian dynamics. Note that the algorithm presented is general enough to handle non-linear non-Gaussian dynamics. In the first example the data shows four trajectories with no clutter, see Figure 1. Figure 2 plots the PHD of position against time. Observe that all four trajectories are automatically initiated and tracked.

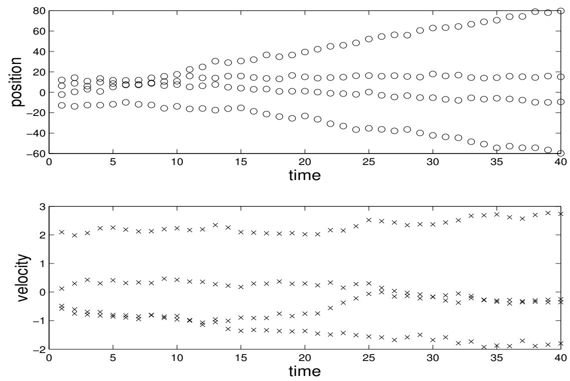

Figure 1: 4 tracks, no death, no birth

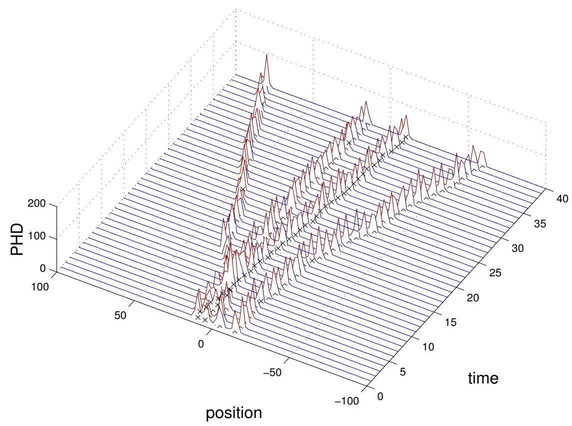

Figure 2: PHD of position

In the second example, we consider an unknown and time varying number of targets observed in clutter. The targets appear spontaneously according to a Poisson process with intensity $0.2N(\cdot \mid 0, 1)$. For simplicity we consider no spawning. 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 3 shows the tracks with clutter on the position measurements and Figure 4 plots the PHD of position against time. Observe from Figure 4 that the PHD filter shows surprisingly good performance, even very short tracks are picked up among clutter.

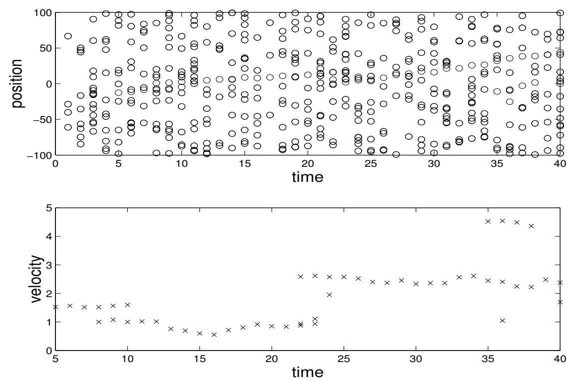

Figure 3: birth rate 0.2, death rate 0.2, clutter rate 10

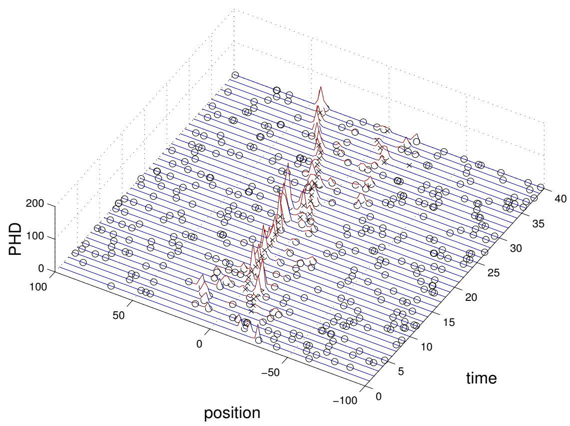

Figure 4: PHD of position

5 Conclusion

In this paper, we have highlighted the relationship between Radon-Nikodym derivative and set derivative of random finite sets. Our main contribution is a sequential Monte Carlo implementation of the probability hypothesis density filter for multi-target tracking. We have demonstrated the efficiency of the algorithm in simulations. There are various potential extensions to this work. First, choosing the importance distributions so as to minimise the (conditional) variance of the weights is a challenging problem. Second, it would be of great practical interest to be able to model targets of different types.

References

[1] A.J. Baddeley and M.N.M. van Lieshout, "ICM for object recognition". In Y. Dodge and J. Whittaker (eds.), *Computational Statistics*, volume 2. Heidelberg: Physica/Springer, pp. 271-286, 1992.

[2] Y. Bar-Shalom and X-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, Storrs, CT: YBS Publishing, 1995.

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

[4] A. Doucet, S. J. Godsill and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comp.*, **10**, 197-208, 2000.

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

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

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

[8] J. Goutsias, R. Mahler and H. Nguyen (eds.), *Random Sets Theory and Applications*, Springer-Verlag New York, 1997.

[9] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models", *J. Comp. Graph. Stat.*, vol. 5, no. 1, pp. 1-25, 1996.

[10] 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.

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

[12] R. Mahler, "A theoretical foundation for the Stein-Winter Probability Hypothesis Density (PHD) multi-target tracking approach," *Proc.2002 MSS Nat'l Symp. on Sensor and Data Fusion*, Vol. I (Unclassified), San Antonio TX, June 2000.

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

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

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