A sequential Monte Carlo method for PHD approximation with conditionally linear/Gaussian models

Mark Morelande
Dept. of Electrical and Electronic Engineering
The University of Melbourne
Parkville, VIC, Australia
mrmore@unimelb.edu.au

Abstract – *A new sequential Monte Carlo procedure for approximating the probability hypothesis density is proposed. The algorithm, based on the replacement of numerical approximation with exact computation, is applicable to the class of conditionally linear/Gaussian models. The proposed algorithm is applied with an efficient, measurement-directed importance density to multiple target tracking using range-bearings measurements. Performance for a given sample size is significantly better than the previously proposed SMC-PHD.*

1 Introduction

The random finite set (RFS) approach to multi-target tracking, introduced by Mahler, is an emerging alternative to the traditional techniques, and has generated substantial research interest [1], [2]. In essence, the collection of target states at any given time is treated as a set-valued multi-target state, and the corresponding collection of sensor measurements is treated as a set-valued multi-target observation. Using RFSs to model the multi-target state and observation, the multi-target tracking problem can be formulated in a Bayesian filtering framework by propagating the posterior density of the multi-target state in time. Due to the inherent combinatorial nature of multi-target densities and the multiple integrations on the (infinite dimensional) multi-target state and observation spaces, the multi-target Bayes recursion is intractable in most practical applications [3, 4].

The probability hypothesis density (PHD) filter [3] was developed as a first moment approximation to the multi-target Bayes recursion. The PHD recursion has the distinct advantage that it operates only on the single-target state space and avoids multiple target data association. A closed form solution for linear Gaussian models was proposed in [5], while a generic sequential Monte Carlo (SMC) implementation with convergence analysis was proposed in [4]. Further convergence results for the SMC-PHD filter are established in [6, 7]. Special cases of this so-called particle-PHD filter have also been independently implemented in [8] and [9]. An auxiliary SMC implementation of the PHD filter, inspired by the auxiliary particle filter (APF) [10], was proposed in [11]. This method boosts the efficiency of the algorithm by pre-selecting particles for propagation according to their relationship with the most recently received observations. A mechanism for clustering particles, required for state estimation, on the basis of the observations to which they are assigned was also proposed. A higher order extension known as the cardinalised PHD filter and closed form solutions have also been developed [12], [13]. Multi-target filters based on the PHD filters have since found successful application in a host of practical problems, see for example [1], [2].

This paper presents a new SMC implementation of the PHD filter which, for a certain class of conditionally linear/Gaussian models, samples only a subset of the elements of the state vector. The contribution to the PHD approximation of the remaining elements of the state vector is computed in closed-form. The exploitation of conditional linearity and Gaussianity in this manner is commonly referred to as Rao-Blackwellisation [14]. Models which satisfy the required conditions are often encountered in tracking applications. An important part of the development of the Rao-Blackwellised SMC-PHD is the use of auxiliary variables, the use of which has also been proposed in [11]. Here an additional auxiliary variable, which determines the nature of target state evolution, is drawn along with the sample and measurement indices. The Rao-Blackwellised SMC-PHD is a general framework which requires the definition of a particular importance density. An importance density is proposed and shown to significantly outperform the SMC-PHD of [4] in a simulation analysis.

The paper is organised as follows. The PHD filter is briefly reviewed in Section 2. The Rao-Blackwellised SMC-PHD is developed in Section 3. A simulation analysis of the proposed algorithm is performed in Section 4.

2 The PHD filter

Let r_t, $t \in \mathbb{N}$ denote the number of targets at time t with states $\boldsymbol{x}_{t,1}, \ldots, \boldsymbol{x}_{t,r_t}$ where $\boldsymbol{x}_{t,i} \in \mathbb{R}^d$. The targets are observed by a sensor which produces a collection of m_t measurements $\boldsymbol{y}_{t,1}, \ldots, \boldsymbol{y}_{t,m_t}$ where $\boldsymbol{y}_{t,j} \in \mathbb{R}^e$. Typically, measurements are due to both targets, although not necessarily all targets, and clutter.

In a random sets framework the target states are collected into the set $X_t = \{\boldsymbol{x}_{t,1}, \ldots, \boldsymbol{x}_{t,r_t}\}$ which is modelled as a realisation of a RFS. Similarly, the measurements are collected into the set $Y_t = \{\boldsymbol{y}_{t,1}, \ldots, \boldsymbol{y}_{t,m_t}\}$, also a realisation of a RFS. The RFS formulation captures important properties of the multiple target tracking problem. First, the unknown number of targets is modelled as a random variable. Similarly, the randomness in the number of measurements is captured by modelling the measurements as a RFS. Second, the use of sets reflects the fact that the ordering of the targets and measurements is usually arbitrary.

The probability hypothesis density (PHD), or intensity, of the multi-target state X_t is the function whose integral over a region $S \subset \mathbb{R}^d$ of the single target state space is equal to the expected number of targets in S [3]. The PHD filter is a recursion for computing the posterior PHD of the multi-target state, i.e., the PHD of X_t conditional on the measurement history $Y_1, \ldots, Y_t$, under certain assumptions [3]. Although the assumptions are not usually met in practice the PHD filter remains a useful method of approximating the PHD. The PHD filter recursion involves several intensities derived from the processes governing the evolution and observation of the multi-target state X_t. A condensed notation is adopted in which $f(\cdot|\boldsymbol{x})$ denotes the intensity of the RFS at time t produced by a target with state $\boldsymbol{x}$ at time $t-1$. This intensity combines the effects of persistent targets, birth targets and splitting targets. The intensity for the measurement RFS produced by a target with state $\boldsymbol{x}$ is $\eta(\boldsymbol{x})g(\cdot|\boldsymbol{x})$ where $\eta(\boldsymbol{x})$ is the detection probability and $g(\cdot|\boldsymbol{x})$ is the measurement PDF. The intensity of the clutter RFS is $\kappa(\cdot)$. Note that the intensities defined above are fixed in time. This is done for the sake of notational brevity and is not required for any of the following analysis.

Let $v_{t-1|t-1}(\cdot)$ denote the posterior PHD of the multi-target state at time $t-1$. Given a set of measurements $\boldsymbol{y}_{t,1}, \ldots, \boldsymbol{y}_{t,m_t}$, the PHD filter computes the posterior PHD at time t in two steps. First, the predicted PHD is computed as

$$v_{t|t-1}(\boldsymbol{x}) = \int f(\boldsymbol{x}|\boldsymbol{z}) v_{t-1|t-1}(\boldsymbol{z})\, d\boldsymbol{z} \qquad (1)$$

Second, the updated PHD is computed as

$$v_{t|t}(\boldsymbol{x}) = \left[1 - \eta + \eta \sum_{j=1}^{m_t} \frac{g(\boldsymbol{y}_{t,j}|\boldsymbol{x})}{\kappa(\boldsymbol{y}_{t,j}) + \eta\, \ell_t(\boldsymbol{y}_{t,j})}\right] v_{t|t-1}(\boldsymbol{x}) \qquad (2)$$

where

$$\ell_t(\boldsymbol{y}) = \int g(\boldsymbol{y}|\boldsymbol{x}) v_{t|t-1}(\boldsymbol{x})\, d\boldsymbol{x} \qquad (3)$$

3 Rao-Blackwellised SMC-PHD

The derivation of the Rao-Blackwellised SMC-PHD (RB-SMC-PHD) uses the following assumptions. The state vector $\boldsymbol{x}_t$ at time t partitions as $\boldsymbol{x}_t = [\boldsymbol{\xi}_t', \boldsymbol{\zeta}_t']'$ such that:

A1 The intensity of the RFS at time t due to a target with state $\boldsymbol{x}_{t-1}$ satisfies

$$\begin{aligned} f(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \sum_{a=1}^{c} \rho_a N(\boldsymbol{\xi}_t - \boldsymbol{\psi}_a(\boldsymbol{\xi}_{t-1}, \boldsymbol{\zeta}_{t-1}), \boldsymbol{Q}_a) \\ \times N(\boldsymbol{\zeta}_t - \boldsymbol{\vartheta}_a(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}, \boldsymbol{\zeta}_{t-1}), \boldsymbol{P}_a) \end{aligned} \qquad (4)$$

where $N(\boldsymbol{\epsilon}, \boldsymbol{\Sigma}) = \exp(-\boldsymbol{\epsilon}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\epsilon}/2)/\sqrt{|2\pi\boldsymbol{\Sigma}|}$ and

$$\boldsymbol{\psi}_a(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \boldsymbol{A}_a\boldsymbol{\xi} + \boldsymbol{B}_a\boldsymbol{\zeta} + \bar{\boldsymbol{\xi}}_a \qquad (5)$$

$$\boldsymbol{\vartheta}_a(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}, \boldsymbol{\zeta}) = \boldsymbol{C}_a\boldsymbol{\xi}_t + \boldsymbol{D}_a\boldsymbol{\xi}_{t-1} + \boldsymbol{E}_a\boldsymbol{\zeta} + \bar{\boldsymbol{\zeta}}_a \qquad (6)$$

A2 The conditional PDF for target-originated measurements satisfies

$$g(\boldsymbol{y}|\boldsymbol{x}) = N(\boldsymbol{y} - \boldsymbol{\phi}(\boldsymbol{\xi}, \boldsymbol{\zeta}), \boldsymbol{R}) \qquad (7)$$

where $\boldsymbol{\phi}(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \boldsymbol{h}(\boldsymbol{\xi}) + \boldsymbol{G}(\boldsymbol{\xi})\boldsymbol{\zeta}$.

A3 The detection probability is assumed to be spatially constant, i.e., $\eta(\boldsymbol{x}) = \eta$.

The portion $\boldsymbol{\xi}_t$ of the state vector is referred to as the nonlinear part of the state vector.

3.1 Development of the RB-SMC-PHD

The following propositions, proofs of which are given in Appendix A, are the basis of the RB-SMC-PHD.

Proposition 1. *If assumption A1 holds and the posterior PHD at time $t-1$ is*

$$v_{t-1|t-1}(\boldsymbol{x}) = \sum_{i=1}^{n} w_{t-1}^i \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_{t-1}^i) N(\boldsymbol{\zeta} - \hat{\boldsymbol{\zeta}}_{t-1}^i, \boldsymbol{\Sigma}_{t-1}^i) \qquad (8)$$

then the prior PHD $v_{t|t-1}(\cdot)$ at time t is

$$\begin{aligned} v_{t|t-1}(\boldsymbol{x}_t) = \sum_{i=1}^{n} w_{t-1}^i \sum_{a=1}^{c} \rho_a N(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{t,a}^i, \boldsymbol{\Psi}_{t,a}) \\ \times N(\boldsymbol{\zeta}_t - \hat{\boldsymbol{\zeta}}_{t|t-1,a}^i(\boldsymbol{\xi}_t), \boldsymbol{\Sigma}_{t|t-1,a}^i) \end{aligned} \qquad (9)$$

where

$$\hat{\boldsymbol{\xi}}_{t,a}^i = \boldsymbol{\psi}_a(\boldsymbol{\xi}_{t-1}^i, \hat{\boldsymbol{\zeta}}_{t-1}^i) \qquad (10)$$

$$\boldsymbol{\Psi}_{t,a}^i = \boldsymbol{B}_a\boldsymbol{\Sigma}_{t-1}^i\boldsymbol{B}_a' + \boldsymbol{Q}_a \qquad (11)$$

$$\hat{\boldsymbol{\zeta}}_{t|t-1,a}^i(\boldsymbol{\xi}_t) = \boldsymbol{\vartheta}_a(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}^i, \hat{\boldsymbol{\zeta}}_{t-1|t,a}^i) \qquad (12)$$

$$\boldsymbol{\Sigma}_{t|t-1,a}^i = \boldsymbol{E}_a\boldsymbol{\Sigma}_{t-1|t,a}^i\boldsymbol{E}_a' + \boldsymbol{P}_a \qquad (13)$$

with

$$\hat{\boldsymbol{\zeta}}^i_{t-1|t,a} = \hat{\boldsymbol{\zeta}}^i_{t-1} + \boldsymbol{W}^i_{t,a}(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}^i_{t,a}) \tag{14}$$

$$\boldsymbol{\Sigma}^i_{t-1|t,a} = \boldsymbol{\Sigma}^i_{t-1} - \boldsymbol{W}^i_{t,a}\boldsymbol{B}_a\boldsymbol{\Sigma}^i_{t-1} \tag{15}$$

where $\boldsymbol{W}^i_{t,a} = \boldsymbol{\Sigma}^i_{t-1}\boldsymbol{B}'_a(\boldsymbol{\Psi}^i_{t,a})^{-1}$.

Proposition 2. *If assumptions A2 and A3 hold and the prior PHD at time t is given by (9), then the posterior PHD* $v_{t|t}(\cdot)$ *at time t is*

$$v_{t|t}(\boldsymbol{x}_t) = \sum_{i=1}^{n}\sum_{a=1}^{c}\sum_{j=0}^{m_t} u^i_{t,a,j}(\boldsymbol{\xi}_t)N(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}^i_{t,a}, \boldsymbol{\Psi}_{t,a}) \times N(\boldsymbol{\zeta}_t - \hat{\boldsymbol{\zeta}}^i_{t|t,a,j}(\boldsymbol{\xi}_t), \boldsymbol{\Sigma}^i_{t|t,a,j}) \tag{16}$$

where

$$u^i_{t,a,j}(\boldsymbol{\xi}_t) = \begin{cases} \rho_a w^i_{t-1}(1-\eta), & j = 0, \\ \rho_a w^i_{t-1}\eta \dfrac{N(\boldsymbol{\epsilon}^i_{t,a,j}(\boldsymbol{\xi}_t), \boldsymbol{S}^i_{t,a})}{\kappa(\boldsymbol{y}_{t,j}) + \eta\ell_t(\boldsymbol{y}_{t,j})}, & j > 0. \end{cases} \tag{17}$$

$$\hat{\boldsymbol{\zeta}}^i_{t|t,a,j}(\boldsymbol{\xi}_t) = \begin{cases} \hat{\boldsymbol{\zeta}}^i_{t|t-1,a}(\boldsymbol{\xi}_t), & j = 0, \\ \hat{\boldsymbol{\zeta}}^i_{t|t-1,a}(\boldsymbol{\xi}_t) + \boldsymbol{K}^i_{t,a}\boldsymbol{\epsilon}^i_{t,a,j}(\boldsymbol{\xi}_t), & j > 0, \end{cases} \tag{18}$$

$$\boldsymbol{\Sigma}^i_{t|t,a} = \begin{cases} \boldsymbol{\Sigma}^i_{t|t-1,a}, & j = 0, \\ (\boldsymbol{I} - \boldsymbol{K}^i_{t,a}\boldsymbol{G}(\boldsymbol{\xi}_t))\boldsymbol{\Sigma}^i_{t|t-1,a}, & j > 0. \end{cases} \tag{19}$$

with

$$\boldsymbol{\epsilon}^i_{t,a,j}(\boldsymbol{\xi}_t) = \boldsymbol{y}_{t,j} - \boldsymbol{\phi}(\boldsymbol{\xi}_t, \hat{\boldsymbol{\zeta}}^i_{t|t-1,a}(\boldsymbol{\xi}_t)) \tag{20}$$

$$\boldsymbol{S}^i_{t,a} = \boldsymbol{G}(\boldsymbol{\xi}_t)\boldsymbol{\Sigma}^i_{t|t-1,a}\boldsymbol{G}(\boldsymbol{\xi}_t)' + \boldsymbol{R} \tag{21}$$

$$\boldsymbol{K}^i_{t,a} = \boldsymbol{\Sigma}^i_{t|t-1,a}\boldsymbol{G}(\boldsymbol{\xi}_t)'(\boldsymbol{S}^i_{t,a})^{-1} \tag{22}$$

Given a sample-based approximation to the posterior PHD $v_{t-1|t-1}(\cdot)$ at time $t-1$, of the form (8), it is desired to produce an approximation to $v_{t|t}(\cdot)$ of the same form. This can be done by drawing from the mixture (16) derived in Proposition 2. It is useful to re-write the mixture (16) as

$$\begin{aligned} v_{t|t}(\boldsymbol{x}_t, i, a, j) &= u^i_{t,a,j}(\boldsymbol{\xi}_t)N(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}^i_{t,a}, \boldsymbol{\Psi}_{t,a}) \\ &\quad \times N(\boldsymbol{\zeta}_t - \hat{\boldsymbol{\zeta}}^i_{t|t,a,j}(\boldsymbol{\xi}_t), \boldsymbol{\Sigma}^i_{t|t,a,j}) \\ &= \omega_{t|t}(\boldsymbol{\xi}_t, i, a, j)\pi^i_{t|t,a,j}(\boldsymbol{\zeta}_t|\boldsymbol{\xi}_t) \end{aligned} \tag{23}$$

Eq. (23) decomposes the updated PHD into the product of an intractable term $\omega_{t|t}(\cdot)$, which is the joint posterior PHD of the nonlinear part $\boldsymbol{\xi}_t$ of the target state and the mixture indices i, a and j and a Gaussian PDF in $\boldsymbol{\zeta}_t$ conditional on $\boldsymbol{\xi}_t, i, a$ and j. The desired sample-based approximation to $v_{t|t}(\cdot)$ can be obtained by sampling $\boldsymbol{\xi}_t$ and the mixture indices i, a and j and then computing $\pi^i_{t|t,a,j}(\cdot|\boldsymbol{\xi}_t)$ for each sample. The mixture indices i, a and j are commonly referred to as auxiliary variables [10].

We concentrate now on obtaining an importance sampling approximation to $\omega_{t|t}(\cdot)$. Let $q(\cdot|\boldsymbol{y}_{t,1:m_t})$ denote an importance density which is non-zero wherever $\omega_{t|t}(\cdot)$ is non-zero. For $s = 1, \ldots, n$, we draw $(\boldsymbol{\xi}^s_t, i^s, a^s, j^s) \sim q(\cdot|\boldsymbol{y}_{t,1:m_t})$ and compute the importance weight

$$\begin{aligned} w^s_t &= \omega_{t|t}(\boldsymbol{\xi}^s_t, i^s, a^s, j^s)/q(\boldsymbol{\xi}^s_t, i^s, a^s, j^s) \\ &= \frac{u^i_{t,a,j}(\boldsymbol{\xi}_t)N(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}^{i^s}_{t,a^s}, \boldsymbol{\Psi}_{t,a})}{q(\boldsymbol{\xi}^s_t, i^s, a^s, j^s)} \end{aligned} \tag{24}$$

The accuracy of the sample-based approximation for a limited sample size depends to a large extent on the importance density. It has been found that producing samples conditional on the current measurement is vital for good performance [15]. An importance density with this property is developed below.

3.2 An efficient importance density

Let $\boldsymbol{\Phi}^i_{t,a} = [\nabla_{\boldsymbol{\xi}}\boldsymbol{\phi}(\boldsymbol{\xi}, \hat{\boldsymbol{\zeta}}^i_{t|t-1,a}(\boldsymbol{\xi}))']'|_{\boldsymbol{\xi}=\hat{\boldsymbol{\xi}}^i_{t,a}}$ denote the gradient of the measurement function. The basis of the importance density is the linearised conditional measurement PDF

$$\tau^i_a(\boldsymbol{y}|\boldsymbol{\xi}) = N(\boldsymbol{y} - \hat{\boldsymbol{\phi}}^i_{t,a}(\boldsymbol{\xi}_t), \boldsymbol{S}^i_{t,a}) \tag{25}$$

where

$$\hat{\boldsymbol{\phi}}^i_{t,a}(\boldsymbol{\xi}_t) = \boldsymbol{\phi}(\hat{\boldsymbol{\xi}}^i_{t,a}, \hat{\boldsymbol{\zeta}}^i_{t|t-1,a}(\hat{\boldsymbol{\xi}}^i_{t,a})) + \boldsymbol{\Phi}^i_{t,a}(\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}^i_{t,a}) \tag{26}$$

is the linearised measurement function. By Lemma 1 (see Appendix A) we can write

$$\tau^i_a(\boldsymbol{y}_{t,j}|\boldsymbol{\xi})N(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}^i_{t,a}, \boldsymbol{\Psi}^i_{t,a}) = \bar{\tau}^i_{a,j}\varpi^i_{a,j}(\boldsymbol{\xi}) \tag{27}$$

where

$$\bar{\tau}^i_{a,j} = N(\boldsymbol{y}_{t,j} - \hat{\boldsymbol{y}}^i_{t,a}, \boldsymbol{U}^i_{t,a}) \tag{28}$$

$$\varpi^i_{a,j}(\boldsymbol{\xi}) = N(\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}^i_{t|t,a,j}, \boldsymbol{\Lambda}^i_{t,a}) \tag{29}$$

with

$$\hat{\boldsymbol{y}}^i_{t,a} = \boldsymbol{\phi}(\hat{\boldsymbol{\xi}}^i_{t,a}, \hat{\boldsymbol{\zeta}}^i_{t|t-1,a}(\hat{\boldsymbol{\xi}}^i_{t,a})) \tag{30}$$

$$\boldsymbol{U}^i_{t,a} = \boldsymbol{\Phi}^i_{t,a}\boldsymbol{\Psi}_{t,a}(\boldsymbol{\Phi}^i_{t,a})' + \boldsymbol{S}^i_{t,a} \tag{31}$$

$$\hat{\boldsymbol{\xi}}^i_{t|t,a,j} = \hat{\boldsymbol{\xi}}^i_{t,a} + \boldsymbol{\Upsilon}^i_{t,a}(\boldsymbol{y}_{t,j} - \hat{\boldsymbol{y}}^i_{t,a}) \tag{32}$$

$$\boldsymbol{\Lambda}^i_{t,a} = \boldsymbol{\Psi}^i_{t,a} - \boldsymbol{\Upsilon}^i_{t,a}\boldsymbol{\Phi}^i_{t,a}\boldsymbol{\Psi}^i_{t,a} \tag{33}$$

where $\boldsymbol{\Upsilon}^i_{t,a} = \boldsymbol{\Psi}^i_{t,a}(\boldsymbol{\Phi}^i_{t,a})'(\boldsymbol{U}^i_{t,a})^{-1}$. The importance density is

$$q(\boldsymbol{\xi}, i, a, j|\boldsymbol{y}_{t,1:m_t}) = \alpha^i_{a,j}\varpi^i_{a,j}(\boldsymbol{\xi})/R \tag{34}$$

where

$$R = \sum_{a,j,i} \alpha^i_{a,j} \tag{35}$$

$$\alpha^i_{a,j} = \begin{cases} \rho_a w^i_{t-1}(1-\eta), & j=0, \\ \rho_a w^i_{t-1}\eta\bar{\tau}^i_{a,j}/[\kappa(\boldsymbol{y}_{t,j}+\eta\ell(\boldsymbol{y}_{t,j})], & j>0. \end{cases} \tag{36}$$

For $j>0$, $\varpi^i_{a,j}(\cdot)$ is given by (29). For $j=0$,

$$\varpi^i_{a,0}(\boldsymbol{\xi}) = N(\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}^i_{t,a}, \boldsymbol{\Psi}^i_{t,a}). \tag{37}$$

The importance weights are, for $s=1,\ldots,n$,

$$w^s_t = \begin{cases} R, & j^s = 0, \\ R\dfrac{N(\boldsymbol{y}_{t,j^s} - \boldsymbol{\phi}(\boldsymbol{\xi}^s_t, \hat{\boldsymbol{\zeta}}^{i^s}_{t|t-1,a^s}(\boldsymbol{\xi}^s_t)), \boldsymbol{S}^{i^s}_{t,a^s})}{\tau^{i^s}_{a^s}(\boldsymbol{y}_{t,j^s}|\boldsymbol{\xi}^s_t)}, & \\ & j^s > 0. \end{cases} \tag{38}$$

A recursion of the SMC-PHD using the importance density (34), and including computation of the conditional posterior PDFs of the elements $\boldsymbol{\zeta}_t$, is summarised in Algorithm 1. The auxiliary variables determine the distribution from which the nonlinear part of the target state is drawn. The auxiliary variable i determines the particle from the sample approximation to $v_{t-1|t-1}(\cdot)$ which will be propagated to time t. Then, conditional on i, the auxiliary variable a determines the manner in which the particle evolves, e.g., persistent, splitting or birth target. Finally, the auxiliary variable j, sampled conditional on the variables i and a, determines which of the measurements, if any, are "associated" with the particle. Note that the mean likelihood $\ell_t(\boldsymbol{y}_{t,j})$ of (3) must be approximated using the sample-based approximation of the posterior PHD at time $t-1$. This has not been shown in the listed algorithm.

Algorithm 1: A recursion of the Rao-Blackwellised auxiliary SMC-PHD

Input: $w^i_{t-1}, \boldsymbol{\xi}^i_{t-1}, \hat{\boldsymbol{\zeta}}^i_{t-1}, \boldsymbol{\Sigma}^i_{t-1}$ for $i=1,\ldots,n$, $\boldsymbol{y}_{t,1},\cdots,\boldsymbol{y}_{t,m_t}$
% *Compute auxiliary variable sampling probabilities*
for $i=1,\ldots,n$ **do**
  set $\beta^i = 0$
  for $a=1,\ldots,c$ **do**
    set $\nu^i_a = 0$
    compute $\hat{\boldsymbol{y}}^i_{t,a}$ and $\boldsymbol{U}^i_{t,a}$ using (30) and (31).
    for $j=0,\ldots,m_t$ **do**
      compute $\alpha^i_{a,j}$ using (36)
      $R \leftarrow R + \alpha^i_{a,j}$
      $\beta^i \leftarrow \beta^i + \alpha^i_{a,j}$
      $\nu^i_a \leftarrow \nu^i_a + \alpha^i_{a,j}$
% *Draw samples*
for $s=1,\ldots,n$ **do**
  set $i^s = k$ with probability β^k/R
  set $a^s = k$ with probability $\nu^{i^s}_k/\beta^{i^s}$
  set $j^s = k$ with probability $\alpha^{i^s}_{a^s,k}/\nu^{i^s}_{a^s}$
  draw $\boldsymbol{\xi}^s_t \sim \varpi^{i^s}_{a^s,j^s}(\cdot)$
  compute w^s_t using (38)
  compute $\hat{\boldsymbol{\zeta}}^i_t = \hat{\boldsymbol{\zeta}}^{i^s}_{t|t,a^s}(\boldsymbol{\xi}^s_t)$ and $\boldsymbol{\Sigma}^i_t = \boldsymbol{\Sigma}^{i^s}_{t|t,a^s}$

4 Performance Analysis

The scenario for the performance analysis involves targets moving in a two-dimensional surveillance region. The targets are observed by a sensor which produces measurements of the targets in polar coordinates. The surveillance region is a 1000×1000 square with the sensor placed at the centre of the square.

The target state is a four-dimensional vector containing the position and velocity in Cartesian coordinates. The intensity $f(\cdot|\cdot)$ of the target RFS at time t is made up of components due to the persistence of the existing target and the birth of new targets. In particular, for a target with state $\boldsymbol{x}_{t-1}$ at time $t-1$,

$$f(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \theta N(\boldsymbol{x}_t; \boldsymbol{F}\boldsymbol{x}_{t-1}, \boldsymbol{Q}) + \sum_{b=1}^{k^2} \varrho_b N(\boldsymbol{x}_t; \bar{\boldsymbol{x}}_b, \boldsymbol{\Omega}_b) \tag{39}$$

The first term in (39) is the intensity of the RFS due to a persistent target with $\theta\in[0,1]$ the persistence probability and

$$\boldsymbol{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \otimes \boldsymbol{I}_2 \tag{40}$$

$$\boldsymbol{Q} = \sigma \begin{bmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{bmatrix} \otimes \boldsymbol{I}_2 \tag{41}$$

with T the sampling period. The process noise intensity is $\sigma = 1/10$. The second term in (39) is the intensity of the RFS due to birth targets. This is a sum of k^2 terms. For the bth term, $\varrho_b > 0$ is the mean number of birth targets with states distributed according to a Gaussian distribution with mean $\bar{\boldsymbol{x}}_b$ and covariance matrix $\boldsymbol{\Omega}_b$. The birth means are arranged on a $k\times k$ grid over the surveillance region with the covariance matrices such that the entire region is covered by the birth RFS. Selection of k involves a trade-off between performance and computational expense. Selecting a small value of k clearly reduces computational expense. However, because larger birth covariance matrices are required as k decreases, sampled values of the target state can be very far from the predicted target state so that the accuracy of the linearised approximation (26) to the measurement function is reduced. The particular parameters used here are $\theta = 0.95$, $T=1$, $\sigma = 1/10$,

$\varrho_b = 1/g^2$ and $k = 6$.

The target state is partitioned so that $\boldsymbol{\xi} = [x, y]'$ contains the position elements and $\boldsymbol{\zeta} = [\dot{x}, \dot{y}]'$ contains the velocity elements. The target measurement PDF satisfies assumption A2 with

$$\boldsymbol{h}(\boldsymbol{\xi}) = \begin{bmatrix} \sqrt{x^2 + y^2} \\ \arctan(y/x) \end{bmatrix} \tag{42}$$

$$\boldsymbol{G}(\boldsymbol{\xi}) = \boldsymbol{0} \tag{43}$$

The target measurement covariance matrix is $\boldsymbol{R} = \mathrm{diag}(16, (\pi/180)^2)$. The detection probability is $\eta = 0.9$. The intensity of the clutter RFS if $\kappa(\boldsymbol{y}) = 10\, U_S(\boldsymbol{y})$ where $S \subset [0, 500\sqrt{2}] \times [-\pi, \pi]$ is the surveillance region in polar space and U_S is the PDF of a uniformly distributed random variable over S.

In the particular scenario considered here the number of targets varies as shown in Figure 1. The initial target positions are uniformly distributed over the surveillance region. The initial target velocities are uniformly distributed over a region such that each has a maximum absolute value of 8 and, in the absence of process noise, the target remains in the surveillance region throughout the surveillance interval. A typical scenario generated under these conditions is shown in Figure 2.

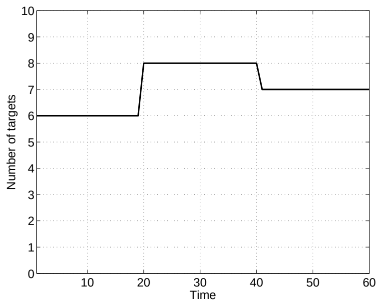

Figure 1: Number of targets in the simulation scenario plotted against time.

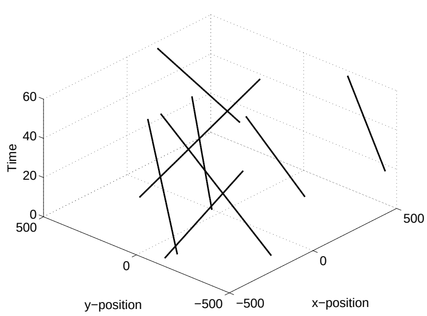

Figure 2: An example realisation of target trajectories for the simulation scenario.

The performance of the proposed algorithm is compared to that of the SMC-PHD of [4]. In the algorithm of [4] specific birth samples are proposed and the number of samples used to approximate the posterior PHD depends on the estimated number of targets. In particular, if n_{t-1} is the number of samples used to approximate the posterior PHD at time $t - 1$, then the prior PHD at time t is represented by $n_{t-1} + b_t$ samples where b_t is the number of birth target samples. The posterior PHD approximation at time t is obtained by calculating the weights of these samples and then resampling according to these weights to produce an approximation of size $n_t = l\hat{r}_t$ samples where l is a pre-defined constant and $\hat{r}_t$ is the estimated number of targets. Note that computational expense is essentially determined by the number $n_{t-1} + b_t$ of samples used to represent the prior PHD since this is the number of particles which must be proposed and weighted. The proposed RB-SMC-PHD does not propose specific birth samples and is implemented with a fixed sample size although it is possible to use a variable sample size.

Algorithm performance is measured using the optimal subpattern assignment (OSPA) distance proposed in [16]. The OSPA distance captures errors in both the estimated number of targets and the estimated target states. Two parameters must be defined to calculate the OPSA distance: the order, which determines sensitivity to outliers, and the cut-off, which determines the relative contributions of errors in target number estimation and target state estimation. In this analysis the order is 2 and the cut-off is 100.

The mean estimated number of targets and the mean OSPA distance, averaged over 100 realisations, are plotted against time in Figure 3 for the SMC-PHD and RB-SMC-PHD with a sample size of 1000. For the SMC-PHD the sample size of 1000 must be divided between persisting and birth particles. We use $n_t = \mathrm{round}(\min\{750, 50\hat{r}_t\})$ persisting particles and $b_t = 1000 - n_{t-1}$ birth particles. As measured by the mean OSPA distance, the RB-SMC-PHD significantly outperforms the SMC-PHD. The plot of target number estimates suggests that the main improvement offered by the RB-SMC-PHD compared to the SMC-PHD is accurate estimation of the number of targets present.

Next, performance as a function of sample size is considered. The mean OSPA distance, averaged over 100 realisations, is found for the SMC-PHD and RB-SMC-PHD using sample sizes between 100 and 10 000. In this comparison, for a total sample size n, the SMC-PHD proposes $n_t = \mathrm{round}(\min\{3n/4, n\hat{r}_t/20\})$ persisting particles and $b_t = n - n_{t-1}$ birth particles. The time-averaged mean OSPA distance is plotted against sample size in Figure 4. A sample size of about 1000 is sufficient for the RB-SMC-PHD to achieve a level of performance, corresponding to a time-averaged mean OSPA distance of 55, at which further increases in the sample size provide no significant improvement. It can be inferred that this is the best performance which can be achieved by a PHD filter. Even with a sample size

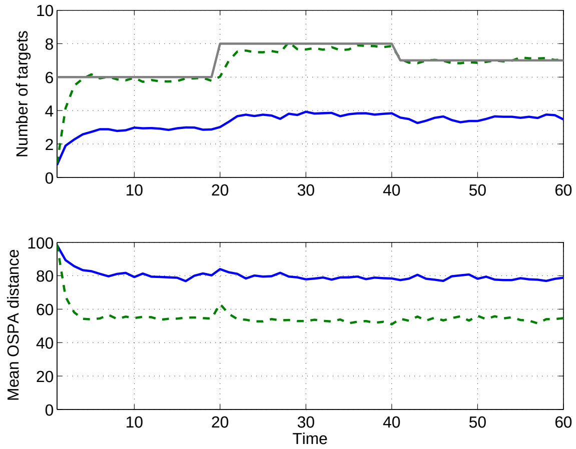

Figure 3: Simulation results for the SMC-PHD (solid blue) and RB-SMC-PHD (dashed green) plotted against time: (top) Mean target number estimate (the solid grey line is the true target number) and (bottom) mean OSPA distance.

of 10 000 the SMC-PHD does not achieve this performance.

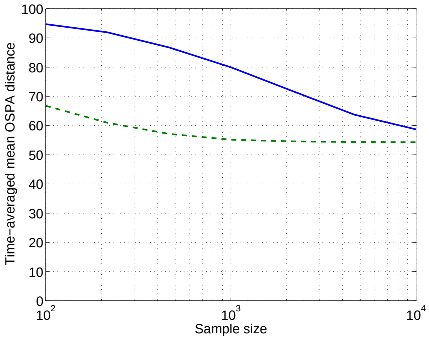

Figure 4: Time-averaged mean OSPA distance plotted against sample size for the SMC-PHD (solid blue) and RB-SMC-PHD (dashed green).

5 Conclusions

The proposed sequential Monte Carlo (SMC) approximation of the probability hypothesis density (PHD) exploits the structure of conditional linear/Gaussian systems to reduce the sample size required for accurate approximation. This was demonstrated for an example involving the tracking of multiple targets using range-bearings measurements.

A number of extensions are possible. A performance comparison with the auxiliary SMC algorithm proposed by Whiteley *et al* would provide a good idea of how much is gained by Rao-Blackwellisation. The proposed procedure should be extended to approximation of the cardinalised PHD which has been demonstrated to produce more reliable estimates of target number than the PHD.

A Proofs

The following lemma is instrumental in the proofs of Proposition 1 and 2 [17]:

Lemma 1. *Assume that $\boldsymbol{z} \in \mathbb{R}^d$, $\boldsymbol{A} \in \mathbb{R}^{d\times m}$, $\boldsymbol{x}, \boldsymbol{\mu} \in \mathbb{R}^m$ and $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$ are positive definite matrices. Then,*

$$N(\boldsymbol{z}-\boldsymbol{A}\boldsymbol{x}, \boldsymbol{B})N(\boldsymbol{x}-\boldsymbol{\mu}, \boldsymbol{\Sigma}) = N(\boldsymbol{z}-\boldsymbol{A}\boldsymbol{\mu}, \boldsymbol{S})N(\boldsymbol{x}-\boldsymbol{\nu}, \boldsymbol{\Omega}), \tag{44}$$

where $\boldsymbol{S} = \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}' + \boldsymbol{B}$ and, with $\boldsymbol{K} = \boldsymbol{\Sigma}\boldsymbol{A}'\boldsymbol{S}^{-1}$,

$$\begin{aligned}\boldsymbol{\nu} &= \boldsymbol{\mu} + \boldsymbol{K}(\boldsymbol{z} - \boldsymbol{A}\boldsymbol{\mu}),\\ \boldsymbol{\Omega} &= \boldsymbol{\Sigma} - \boldsymbol{K}\boldsymbol{A}\boldsymbol{\Sigma}.\end{aligned}$$

Proposition 1 is proved first. Under the conditions of the proposition

$$v_{t-1|t-1}(\boldsymbol{x}) = \sum_{i=1}^{n} w_{t-1}^i \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_{t-1}^i) N(\boldsymbol{\zeta} - \hat{\boldsymbol{\zeta}}_{t-1}^i, \boldsymbol{\Sigma}_{t-1}^i) \tag{45}$$

Substituting (4) (assumption A1) and (45) into the prediction equation for the PHD filter (1) and using Lemma 1 gives

$$\begin{aligned}v_{t|t-1}(\boldsymbol{x}_t) = \sum_{i=1}^{n} w_{t-1}^i \sum_{a=1}^{c} \rho_a N(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{t,a}^i, \boldsymbol{\Psi}_{t,a})\\ \times \int N(\boldsymbol{\zeta}_t - \boldsymbol{\vartheta}_a(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}^i), \boldsymbol{\zeta}_{t-1}, \boldsymbol{P}_a)\\ \times N(\boldsymbol{\zeta}_{t-1} - \hat{\boldsymbol{\zeta}}_{t-1|t,a}^i, \boldsymbol{\Sigma}_{t-1|t,a}^i)\, d\boldsymbol{\zeta}_{t-1}\end{aligned} \tag{46}$$

where

$$\hat{\boldsymbol{\xi}}_{t,a}^i = \boldsymbol{\psi}_a(\boldsymbol{\xi}_{t-1}^i, \hat{\boldsymbol{\zeta}}_{t-1}^i) \tag{47}$$

$$\boldsymbol{\Psi}_{t,a}^i = \boldsymbol{B}_a \boldsymbol{\Sigma}_{t-1}^i \boldsymbol{B}_a' + \boldsymbol{Q}_a \tag{48}$$

$$\hat{\boldsymbol{\zeta}}_{t-1|t,a}^i = \hat{\boldsymbol{\zeta}}_{t-1}^i + \boldsymbol{W}_{t,a}^i(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{t,a}^i) \tag{49}$$

$$\boldsymbol{\Sigma}_{t-1|t,a}^i = \boldsymbol{\Sigma}_{t-1}^i - \boldsymbol{W}_{t,a}^i \boldsymbol{B}_a \boldsymbol{\Sigma}_{t-1}^i \tag{50}$$

with $\boldsymbol{W}_{t,a}^i = \boldsymbol{\Sigma}_{t-1}^i \boldsymbol{B}_a'(\boldsymbol{\Psi}_{t,a}^i)^{-1}$. Applying Lemma 1 to the integrand of (46) gives the desired result.

Proposition 2 is now proved. Substituting (7) (assumption A2) and (9) into the PHD filter update equation (2) gives

$$\begin{aligned}v_{t|t}(\boldsymbol{x}_t) = \sum_{i=1}^{n} w_{t-1}^i \sum_{a=1}^{c} \rho_a N(\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{t,a}^i, \boldsymbol{\Psi}_{t,a})\\ \times N(\boldsymbol{\zeta}_t; \hat{\boldsymbol{\zeta}}_{t|t-1,a}^i(\boldsymbol{\xi}_t), \boldsymbol{\Sigma}_{t|t-1,a}^i)\\ \times \left[1 - \eta + \eta \sum_{j=1}^{m_t} \frac{N(\boldsymbol{y}_{t,j} - \boldsymbol{\phi}(\boldsymbol{\xi}_t, \boldsymbol{\zeta}_t), \boldsymbol{R})}{\kappa(\boldsymbol{y}_{t,j}) + \eta\, \ell_t(\boldsymbol{y}_{t,j})}\right]\end{aligned} \tag{51}$$

Note that assumption A3, which supposes a spatially constant detection probability, has been applied. Applying Lemma 1 to the product

$$N(\boldsymbol{y}_{t,j} - \boldsymbol{\phi}(\boldsymbol{\xi}_t, \boldsymbol{\zeta}_t), \boldsymbol{R}) N(\boldsymbol{\zeta}_t - \hat{\boldsymbol{\zeta}}^i_{t|t-1,a}(\boldsymbol{\xi}_t), \boldsymbol{\Sigma}^i_{t|t-1,a})$$

gives the desired result.

References

[1] R. Mahler, *Statistical Multisource Multitarget Information Fusion.* Artech House, 2007.

[2] B.-T. Vo, "Random finite sets in multi-object filtering," Ph.D. dissertation, University of Western Australia, 2008.

[3] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[4] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multi-target filtering with random finite sets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, 2005.

[5] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[6] A. Johansen, S. Singh, A. Doucet, and B.-N. Vo, "Convergence of the SMC implementation of the PHD filter," *Methodology and Computing in Applied Probability*, vol. 8, no. 2, pp. 265–291, 2006.

[7] D. Clark and J. Bell, "Convergence results for the particle PHD filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2652–2661, 2006.

[8] H. Sidenbladh, "Multi-target particle filtering for the probability hypothesis density," in *Proceedings of the International Conference on Information Fusion*, Cairns, Australia, 2003, pp. 800–806.

[9] T. Zajic and R. Mahler, "A particle-systems implementation of the PHD multi-target tracking filter," in *Proceedings of the SPIE*, vol. 5096, 2003, pp. 291–299.

[10] M. Pitt and N. Shephard, "Filtering via simulation: auxiliary particle filters," *Journal of the American Statistical Association*, vol. 94, pp. 590–599, 1999.

[11] N. Whiteley, S. Singh, and S. Godsill, "Auxiliary particle implementation of the probability hypothesis density filter," in *Proceedings of the IEEE International Symposium on Image and Signal Processing Analysis*, Istanbul, Turkey, 2007.

[12] R. Mahler, "PHD filters of higher order in target number," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 1523–1543, 2007.

[13] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3553–3567, 2007.

[14] G. Casella and C. Robert, "Rao-Blackwellisation of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 81–94, 1996.

[15] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[16] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3447–3457, 2008.

[17] Y. Ho and R. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Transactions on Automatic Control*, vol. 9, pp. 333–339, 1964.