

# Rao-Blackwellized Particle Filter for Multiple Target Tracking

Simo Särkkä\*, Aki Vehtari, Jouko Lampinen

*Helsinki University of Technology, Finland*

---

## Abstract

In this article we propose a new Rao-Blackwellized particle filtering based algorithm for tracking an unknown number of targets. The algorithm is based on formulating probabilistic stochastic process models for target states, data associations, and birth and death processes. The tracking of these stochastic processes is implemented using sequential Monte Carlo sampling or particle filtering, and the efficiency of the Monte Carlo sampling is improved by using Rao-Blackwellization.

*Key words:* multiple target tracking, data association, unknown number of targets, Rao-Blackwellization, sequential Monte Carlo, particle filtering

---

---

<sup>1</sup> Simo Särkkä\* (the corresponding author) is with Helsinki University of Technology, P.O. Box 9203. FIN-02015 HUT, Finland. E-mail: simo.sarkka@hut.fi, Tel: +358 40 757 0730, Fax: +358 9 451 4830.

<sup>2</sup> Aki Vehtari is with Helsinki University of Technology, P.O. Box 9203. FIN-02015 HUT, Finland. E-mail: aki.vehtari@hut.fi, Tel: +358 9 451 4849, Fax: +358 9 451 4830.

<sup>3</sup> Jouko Lampinen is with Helsinki University of Technology, P.O. Box 9203. FIN-02015 HUT, Finland. E-mail: jouko.lampinen@hut.fi, Tel: +358 9 451 4827, Fax: +358 9 451 4830.

## 1 Introduction

This article is an extended version<sup>4</sup> of the article [1], in which we proposed a *Rao-Blackwellized particle filtering* based multiple target tracking algorithm called *Rao-Blackwellized Monte Carlo data association* (RBMCDCA). In this article we extend the RBMCDCA algorithm to tracking *an unknown and time varying number of targets*. In the proposed algorithm, not only the target states and the data associations, but also the births and deaths of the targets are modeled as hidden stochastic processes, which are observed through the measurements. The states of the stochastic processes are inferred from the measurements using Rao-Blackwellized particle filtering.

In principle, it is always possible to approximate the optimal Bayesian filtering solution [2] to the estimation problem with a particle filter [3–5], which generates samples from the joint posterior distribution of the hidden stochastic processes. The idea behind Rao-Blackwellization is that by conditioning on the data associations and the birth and death processes, the posterior distributions of the states of the targets can be approximated with Gaussian distributions. For this reason, the target states can be integrated out analytically and the particle filter only needs to be applied to the data associations and the birth and death processes. This significantly reduces the computational requirements and increases the efficiency of the particle filter.

The problem of tracking an unknown number of targets can be divided into the subproblems of *tracking single targets*, *data association* and *estimating the number of targets*. Of course, these subproblems are coupled such that one cannot be solved without the other, but this conceptual division is often done in literature. In the article [1] we presented a solution to first two of these problems using Rao-Blackwellized particle filtering together with classical filtering theory [2,6]. The main contribution of this article is to solve the problem of estimating the number of targets.

In the next two sections we shall present a short review of the existing methods for data association in the case of a known number of targets (Section 1.1) and for data association and number of targets estimation in the case of an unknown number of targets (Section 1.2). Relationships with the method described in this article are also discussed.

---

<sup>4</sup> A short version of this study was presented at the Seventh International Conference on Information Fusion (FUSION 2004) in Stockholm, Sweden [1]

### 1.1 Approaches to Data Association

The classical data association methods for multiple target tracking can be divided into two main classes [7]. *Unique-neighbor data association* methods, such as *multiple hypothesis tracking* (MHT), associate each measurement with one of the previously established tracks. *All-neighbors data association* methods, such as *joint probabilistic data association* (JPDA), use all measurements for updating all the track estimates.

The idea of MHT [8,7,9] is to associate each measurement with one of the existing tracks, or to form a new track from the measurement. Because this association is not necessarily unique, several hypotheses are continuously formed and maintained. The MHT algorithm calculates the likelihoods of the measurements and the posterior probabilities of the hypotheses, storing only the most probable hypotheses. To enhance the computational efficiency, heuristic methods such as gating, hypothesis merging, clustering and several other strategies can be employed.

*Probabilistic multiple hypothesis tracking* (PMHT) [10] is a modification of the MHT, where the data associations are assumed to be independent over the target tracks. This way the computational complexity of the method is substantially reduced, but it is also impossible to model certain practical constraints, for example, to restrict the number of contacts per target to one on each scan. Our RBMCDA method [1] in its original form had this same restriction, but it can be easily overcome by allowing dependencies in the data association priors, as shown in Section 2.3.

JPDA [8,7] approximates the posterior distributions of the targets as separate Gaussian distributions for each target. If the number of targets is  $T$ , then  $T$  separate Gaussian distributions are maintained. The number of Gaussian distributions is kept constant by integrating over the distribution of data associations of the previous step. This results in an algorithm where each of the target estimates gets updated by every measurement with weights that depend on the predicted probabilities of the associations. Gating is used for limiting the number of measurements for each track. If the predicted probabilities are too low (i.e., below a predefined threshold) for certain targets, those targets are not updated at all. Clutter measurements can be modeled similarly.

*Sequential Monte Carlo* (SMC) based multiple target tracking methods [11–13] typically belong to the class of *unique-neighbor data association* methods, as they are based on representing the data association and state posteriors as discrete sets of hypotheses. These kind of SMC methods can be considered generalizations of MHT. Instead of maintaining the  $N$  most probable data association hypotheses, the joint tracking and data association problem

is modeled as a Bayesian estimation problem and the *posterior distribution* is estimated with SMC methods. This *particle filtering* approach has the advantage that there are no restrictions on the analytic form of model, although the required number of particles for a given accuracy can be high.

In article [1] we proposed how SMC based tracking and data association algorithms can be made more accurate and efficient using Rao-Blackwellization. In the Rao-Blackwellized Monte Carlo data association (RBMCD) algorithm the states are integrated out in closed form and SMC is only used for the data association indicators. Instead of a pure particle representation, this leads to a mixture of Gaussians representation of the joint posterior distribution, which reduces variance and requires less particles for the same accuracy.

## 1.2 Approaches to Tracking Unknown Number of Targets

The JPDA method was originally formulated for a known number of targets, but it is possible to include track formation and termination logic in cascade with the algorithm [8]. There is no explicit probability model for target appearance and disappearance, but instead there can be a Markov chain model for the number of data associations before track initiation is confirmed. Tracks are terminated when the probability of target existence goes below a predefined threshold.

MHT based methods [8,7,9] as well as PMHT [10] form hypotheses of associations with new targets for every measurement. In practice, to reduce the computational complexity, new target hypotheses (or new track hypotheses) are formed only when the measurement falls into an area where the likelihood of the association with the existing targets is too low. A track is deleted when its likelihood becomes too low compared to the other tracks.

Random sets and finite set statistics (FISST) [14] provide a very general framework for Bayesian modeling of multiple target tracking in the case of an unknown number of targets. A tractable implementation of the framework is to use the first order moment of the multi-target posterior, the probability hypothesis density (PHD) [15] as an approximation. SMC based implementations of the PHD have been reported, for example, in the articles [16] [17].

In the SMC based method presented in the article [13] the extension to an unknown number of targets is based on hypothesis testing. Because the algorithm generates estimates of data association probabilities, these estimates can be used for approximating the probability of the hypothesis that the target has disappeared from the surveillance area. The detection of the appearance of a new target is based on testing the hypothesis between association with the old targets and with the new target.

The article [18] presents a SMC based method, which is similar to our method except that a plain particle presentation of the joint posterior distribution is used. In the method, birth and death moves in particle proposals are used, and the moves resemble the birth and death models used in this article. Due to the plain particle presentation, the method in the article [18] is also applicable to the more general case of target tracking without explicit thresholding of measurements.

The method in [19] also resembles our method, except that the article does not suggest any particular form for the birth and death models. The approximation based on limiting the number of births and deaths on each time step is also discussed in [19].

The particle filtering based method in [20] uses exponential (Poisson) models for target appearance and disappearance a bit similarly to our method. The branching particle based solution [21] also models target appearance as a stochastic (Markov) process.

The tracking of an unknown number of targets is also closely related to model selection. An application of SMC methods to estimating the number of RBF network weights from data is presented in [22].

In this article, we extend the SMC based RBMCDA method [1] to tracking an unknown number of targets. The extension is based on modeling the birth and death stochastic processes, such that track formation and termination (or initiation and deletion) are not based on heuristic rules, but on the rules determined by the estimation algorithm designed for the probability model. However, this probability model of extension to an unknown number of targets is closer to the approach presented in [9] than to the approach in [15] despite the more strict Bayesian nature of the latter.

### 1.3 Particle Filtering

Sequential Importance Resampling (SIR) (see, e.g. [3–5,23]), is a generalization of the *particle filtering* framework for the estimation of state space models of the form

$$\begin{aligned}\mathbf{x}_k &\sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) \\ \mathbf{y}_k &\sim p(\mathbf{y}_k \mid \mathbf{x}_k),\end{aligned}\tag{1}$$

where  $\mathbf{x}_k \in \mathbb{R}^n$  is the state at time step  $t_k$  and  $\mathbf{y}_k \in \mathbb{R}^d$  is the measurement. The algorithm uses a weighted Monte Carlo representation of the posterior state distribution. This set of particles is updated and reweighted using a recursive version of importance sampling. An additional *resampling* procedure

is used for removing particles with very small weights and duplicating particles with large weights. The variance introduced by the resampling procedure can be reduced by proper choice of the resampling method [4].

In our version of SIR resampling is not performed on every step, but only when it is actually needed. One way of implementing this is to do resampling on every  $k$ th step, where  $k$  is some predefined constant. This method has the advantage that it is unbiased. Another way, which is also used in our simulation system, is *adaptive resampling*, in which the effective number of particles, which is estimated from the variance of the particle weights [24] is used for monitoring the need for resampling.

The performance of the SIR algorithm is also dependent on the importance distribution  $\pi(\cdot)$ , which is an approximation of the posterior distribution of states given the values at the previous step. The importance distribution should be of such functional form that we can easily draw samples from it and evaluate the probability densities of the sample points. *The optimal importance distribution* (see, e.g., [5,23]) is  $\pi(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{1:k}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{1:k})$ . This importance distribution is optimal in the sense that it minimizes the variance of the importance weights.

One way of improving the efficiency of SMC is to use Rao-Blackwellization. The idea of the *Rao-Blackwellized particle filter* (RBPF) [25,5,23] is that sometimes it is possible to evaluate some of the filtering equations analytically and the others with Monte Carlo sampling instead of computing everything with pure sampling. According to the *Rao-Blackwell theorem* this leads to estimators with less variance than what could be obtained with pure Monte Carlo sampling [26]. An intuitive way of understanding this is that the marginalization replaces the finite Monte Carlo particle set representation with an infinite closed form particle set, which is always more accurate than any finite set.

The sampling and resampling approach that we used is not necessarily the most efficient in all conditions, but it turned out to work well in our applications. By tuning the resampling algorithm and possibly changing the order of weight computation and sampling, accuracy and computational efficiency of the algorithm could possibly be improved [27]. An important issue is that sampling could be more efficient without replacement, such that duplicate samples are not stored. There is also evidence that in some situations it is more efficient to use a simple deterministic algorithm for preserving the  $N$  most likely particles. In the article [28] it is shown that in digital demodulation, where the sampled space is discrete and the optimization criterion is the minimum error, the deterministic algorithm performs better.

## 2 RBMCDA with a Known Number of Targets

This section reviews the Rao-Blackwellized Monte Carlo data association (RBM-CDA) method that we proposed in [1] and relates it to the Rao-Blackwellized particle filtering framework. The method described here is basically the same as in the original article except that the requirement of IID prior data association probabilities is replaced with a Markov chain assumption.

### 2.1 Filtering Model

The filtering model of the RBMCDA algorithm is the following:

- *Clutter or false alarm measurements* can be modeled using any probability density, which is independent of the target states  $\mathbf{x}_k = (\mathbf{x}_{k,1} \cdots \mathbf{x}_{k,T})^T$

$$p(\mathbf{y}_k \mid c_k = 0). \quad (2)$$

Here  $c_k$  is the data association indicator, which has the value  $c_k = 0$  for clutter and  $c_k = j$  for the targets  $j = 1, \dots, T$ . The clutter measurements can be, for example, uniformly distributed in the measurement space of volume  $V$

$$p(\mathbf{y}_k \mid c_k = 0) = 1/V. \quad (3)$$

- *Target measurements* are linear Gaussian

$$p(\mathbf{y}_k \mid \mathbf{x}_{k,j}, c_k = j) = N(\mathbf{y}_k \mid \mathbf{H}_{k,j}\mathbf{x}_{k,j}, \mathbf{R}_{k,j}). \quad (4)$$

where the measurement matrices  $\mathbf{H}_{k,j}$  and covariance matrices  $\mathbf{R}_{k,j}$  can be different for each target. Non-linear measurement models can be used by replacing the non-linear model with a *locally linearized* model as in the extended Kalman filter (EKF) [2,6] or by using the *unscented transformation* as in the unscented Kalman filter (UKF) [29].

- *Target dynamics* are linear Gaussian

$$p(\mathbf{x}_{k,j} \mid \mathbf{x}_{k-1,j}) = N(\mathbf{x}_{k,j} \mid \mathbf{A}_{k-1,j}\mathbf{x}_{k-1,j}, \mathbf{Q}_{k-1,j}), \quad (5)$$

where the transition matrix  $\mathbf{A}_{k-1,j}$  and process noise covariance matrix  $\mathbf{Q}_{k-1,j}$  may be different for different targets. The motions of individual targets are a priori independent. Because for any continuous time linear Gaussian dynamic model there exists an equivalent discrete time linear Gaussian model [2,6], this directly generalizes to continuous time linear models. Non-linear discrete or continuous time dynamic models can be also used as in EKF [2,6] or UKF [29]. With slight modifications to the algorithm, it would be possible to use the interacting multiple model (IMM) filter [6], which allows the modeling of target maneuvers.

- *Target and clutter association priors* are known and can be modeled as an  $m$ th order Markov chain

$$p(c_k \mid c_{k-1}, \dots, c_{k-m}). \quad (6)$$

This kind of higher order models are needed for implementing constraints to the data associations on a single measurement scan (see Section 2.3). Of course, the model can be also a first order Markov chain ( $m = 1$ ) or the associations can be completely independent ( $m = 0$ ).

- *The number of targets  $T$*  is known and constant.
- *Target state priors* can be represented as a weighted importance sample set

$$p(\mathbf{x}_{0,j}) = \sum_i w^{(i)} \mathbf{N}(\mathbf{x}_{0,j} \mid \mathbf{m}_{0,j}^{(i)}, \mathbf{P}_{0,j}^{(i)}). \quad (7)$$

## 2.2 Relationship with the Rao-Blackwellized Particle Filter

The RBMCDA model described in the previous section fits exactly to the Rao-Blackwellized particle filtering framework (see, e.g., [5]) when the sampled latent variable  $\boldsymbol{\lambda}_k$  is defined to contain the data association event indicators,  $c_k$  at time step  $k$

$$\boldsymbol{\lambda}_k = c_k. \quad (8)$$

- (1) The state  $\mathbf{x}_k$  on time step  $k$  consists of the stacked vector of target states

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_{k,1} \\ \vdots \\ \mathbf{x}_{k,T} \end{pmatrix}. \quad (9)$$

- (2) The prior distribution of the joint state  $p(\mathbf{x}_0)$  is Gaussian, because the individual target prior distributions are Gaussian.
- (3) The joint dynamic model of targets is linear Gaussian

$$p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) = \mathbf{N}(\mathbf{x}_k \mid \mathbf{A}_{k-1}\mathbf{x}_{k-1}, \mathbf{Q}_{k-1}), \quad (10)$$

where  $\mathbf{A}_{k-1}$  is a block diagonal matrix consisting of the dynamic models of the targets and  $\mathbf{Q}_{k-1}$  is the block diagonal process noise covariance.

- (4) The joint measurement model of the targets can be written as

$$p(\mathbf{y}_k \mid \mathbf{x}_k, c_k) = \mathbf{N}(\mathbf{y}_k \mid \mathbf{H}_k(c_k)\mathbf{x}_k, \mathbf{R}_k(c_k)), \quad (11)$$

where the measurement model matrix  $\mathbf{H}_k(c_k)$  is formed conditional on the data association  $c_k$  such that the only nonzero entries are the ones



corresponding to the measurement model of the target  $c_k$ . The measurement noise covariance  $\mathbf{R}_k(c_k)$  is the measurement noise covariance of the target  $c_k$ .

- (5) The clutter measurements are simply state independent measurements with a measurement model of the form  $p(\mathbf{y}_k \mid c_k = 0)$ .
- (6) The data associations are modeled as a Markov chain of latent variables as in the Rao-Blackwellized particle filter model.

### 2.3 Sequential Measurement Update Issues

In Section 2.1 we formulated the model such that exactly one measurement is obtained at one time step  $t_k$ . This way, the data association procedure can be reduced to processing one measurement at a time. However, this does not mean that we are restricted to one measurement per time instance (i.e., scan), because the successive time steps may occur on the same time instance.

scan 1 (3 measurements):	$t_1 = t_2 = t_3 = 1$ s
scan 2 (2 measurements):	$t_4 = t_5 = 2$ s
scan 3 (3 measurements):	$t_6 = t_7 = t_8 = 2.5$ s
scan 4 (1 measurement):	$t_9 = 3.5$ s
...	...

Table 1

Several measurement steps can occur on the same time instance or scan.

In the example presented in Table 1 the time steps 1,2 and 3 all occur on time instance 1 s, time steps 4 and 5 occur on time instance 2 s and so on. Now, for example, time steps 1,2 and 3 can be processed sequentially by setting  $\Delta t = 0$  in the dynamic model and in the prediction step, which is equivalent to not performing the prediction step at all between the measurements.

The original RBMCDA that was presented in [1] had the restriction that the data associations were required to be conditionally independent as in PMHT [10]. This makes it impossible to model certain joint effects in sets of measurements obtained on one time instance (i.e., scan). For example, it is impossible to restrict the maximum number of data associations with each target to one per time instance. However, by allowing the data association priors to depend on previous data associations (at least on the same time instance) this restriction can be included in the model.

As discussed in Section 2.1, the data association indicators may also depend on previous data associations, that is, they may form an  $m$ th order Markov chain

$$p(c_k \mid c_{k-1}, \dots, c_{k-m}). \quad (12)$$

This kind of model can be used for restricting the data associations to at most one data association to each target as follows:

- The joint prior model can be written in the general form

$$p(c_{k+m-1}, \dots, c_k). \quad (13)$$

We can encode the desired restriction into the prior by assigning zero probability to any joint event, which would have two associations to the same target.

- The joint model can be also expanded as follows:

$$\begin{aligned} p(c_{k+m-1}, \dots, c_k) \\ = \prod_{j=1}^m p(c_{k+j} \mid c_k, \dots, c_{k+j-1}). \end{aligned} \quad (14)$$

This means that exactly the same model can be realized by using the following priors for the data associations:

- $c_k$  has the prior  $p(c_k)$ .
- $c_{k+1}$  has the prior  $p(c_{k+1} \mid c_k)$ .
- ...
- $c_{k+m-1}$  has the prior  $p(c_{k+m-1} \mid c_{k+m-2}, \dots, c_k)$ .

Using the above idea it is possible to reduce any joint data association prior of form (13) to an equivalent Markov prior model, which is suitable for sequential processing.

Assume, for example, that we are tracking two targets and on each time instance we may obtain zero or one detections from each of the targets. The rest of the detections are false alarms, that is, clutter. The target detection probabilities of both targets are the same and given as

$$p(\text{detection}) = p_d. \quad (15)$$

Assume that on time instance  $t_{\text{scan}}$  we obtain  $m$  measurements  $\mathbf{y}_k, \dots, \mathbf{y}_{k+m-1}$  (i.e., the times of the steps are  $t_k, \dots, t_{k+m-1} = t_{\text{scan}}$ ). The prior for the data associations can be now defined sequentially as follows:

- Define detection indicators as follows

$$\begin{aligned} \delta_1(j) &= \begin{cases} 1, & \text{if there is target 1 detection in } c_{k+j-1} \dots c_k \\ 0, & \text{otherwise} \end{cases} \\ \delta_2(j) &= \begin{cases} 1, & \text{if there is target 2 detection in } c_{k+j-1} \dots c_k \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (16)$$

- Compute the data association priors given each of the possible detection indicator combinations:

$$\begin{aligned}
& p(c_{k+j} \mid \delta_1(j) = 0 \text{ and } \delta_2(j) = 0) \\
& p(c_{k+j} \mid \delta_1(j) = 0 \text{ and } \delta_2(j) = 1) \\
& p(c_{k+j} \mid \delta_1(j) = 1 \text{ and } \delta_2(j) = 0) \\
& p(c_{k+j} \mid \delta_1(j) = 1 \text{ and } \delta_2(j) = 1).
\end{aligned} \tag{17}$$

Prior models having the restriction of one data association to each target per time instance can be defined in similar manner for any number of targets. The resulting prior is of the recursive (Markov) form, which is a special case of the model described in Section 2.1.

#### 2.4 Data Representation

The algorithm state consists of a set of  $N$  particles, where each particle  $i$  at time step  $k$  contains the following:

$$\{c_{k-m+1:k}^{(i)}, \mathbf{m}_{k,1}^{(i)}, \dots, \mathbf{m}_{k,j}^{(i)}, \dots, \mathbf{m}_{k,T}^{(i)}, \mathbf{P}_{k,1}^{(i)}, \dots, \mathbf{P}_{k,j}^{(i)}, \dots, \mathbf{P}_{k,T}^{(i)}, w_k^{(i)}\}, \tag{18}$$

where

- $c_{k-m+1:k}^{(i)}$  are the data association indicators of time steps  $k - m + 1, \dots, k$  with integer values  $0, \dots, T$ , where  $T$  is the number of targets. If the data association prior model is an  $m$ th order Markov model, then  $m$  previous data associations should be stored. If the data association prior is time independent, the data association indicators do not need to be stored at all.
- $\mathbf{m}_{k,j}^{(i)}, \mathbf{P}_{k,j}^{(i)}$  are the mean and covariance of the target  $j$ , and they are conditional on the data association history  $c_{1:k}^{(i)}$ .
- $w_k^{(i)}$  is the importance weight of the particle.

#### 2.5 Evaluating and Sampling from the Optimal Importance Distribution

For each particle  $i$ , the optimal importance distribution is given by

$$p(c_k \mid \mathbf{y}_{1:k}, c_{1:k-1}^{(i)}). \tag{19}$$

The marginal measurement likelihood is given by [1]

$$p(\mathbf{y}_k \mid c_k, \mathbf{y}_{1:k-1}, c_{1:k-1}^{(i)}) = \begin{cases} 1/V & \text{if } c_k = 0 \\ \text{KF}_{lh}(\mathbf{y}_k, \mathbf{m}_{j,k}^{-(i)}, \mathbf{P}_{j,k}^{-(i)}, \mathbf{H}_{j,k}, \mathbf{R}_{j,k}) & \text{if } c_k = j \end{cases} \quad (20)$$

where  $j = 1, \dots, T$  and  $\text{KF}_{lh}(\cdot)$  denotes the Kalman filter measurement likelihood evaluation.  $\mathbf{H}_{j,k}$  and  $\mathbf{R}_{j,k}$  are the measurement model matrix and the measurement covariance matrix of the target  $j$ , respectively. For  $j = 1, \dots, T$  we have

$$[\mathbf{m}_{j,k}^{-(i)}, \mathbf{P}_{j,k}^{-(i)}] = \text{KF}_p(\mathbf{m}_{j,k-1}^{(i)}, \mathbf{P}_{j,k-1}^{(i)}, \mathbf{A}_{j,k-1}, \mathbf{Q}_{j,k-1}), \quad (21)$$

where  $\text{KF}_p(\cdot)$  denotes the Kalman filter prediction step, and  $\mathbf{m}_{j,k-1}^{(i)}, \mathbf{P}_{j,k-1}^{(i)}$  are the mean and the covariance of target  $j$  in particle  $i$ , which is conditioned on the state history  $c_{1:k-1}^{(i)}$ .  $\mathbf{A}_{j,k-1}$  and  $\mathbf{Q}_{j,k-1}$  are the transition matrix of dynamic model and the process noise covariance matrix of the target  $j$ , respectively.

The posterior distribution of  $c_k$  can be calculated using Bayes' rule

$$p(c_k \mid \mathbf{y}_{1:k}, c_{1:k-1}^{(i)}) \propto p(\mathbf{y}_k \mid c_k, \mathbf{y}_{1:k-1}, c_{1:k-1}^{(i)}) \times p(c_k \mid c_{k-m:k-1}^{(i)}), \quad (22)$$

where we have used the fact that an association  $c_k$  does not depend on the previous measurements  $\mathbf{y}_{1:k-1}$ , and depends only on the  $m$  previous associations  $c_{k-m:k-1}$  if the order of the Markov model is  $m$ .

We can sample from the optimal importance distribution as follows:

- (1) Compute the unnormalized clutter association probability

$$\hat{\pi}_0^{(i)} = p(\mathbf{y}_k \mid c_k^{(i)} = 0, \mathbf{y}_{1:k-1}, c_{1:k-1}^{(i)}) \times p(c_k^{(i)} = 0 \mid c_{k-m:k-1}^{(i)}). \quad (23)$$

- (2) Compute the unnormalized target association probabilities for each target  $j = 1, \dots, T$

$$\hat{\pi}_j^{(i)} = p(\mathbf{y}_k \mid c_k^{(i)} = j, \mathbf{y}_{1:k-1}, c_{1:k-1}^{(i)}) \times p(c_k^{(i)} = j \mid c_{k-m:k-1}^{(i)}). \quad (24)$$

- (3) Normalize the importance distribution:

$$\pi_j^{(i)} = \frac{\hat{\pi}_j^{(i)}}{\sum_{j'=0}^T \hat{\pi}_{j'}^{(i)}}, \quad j = 0, \dots, T. \quad (25)$$

- (4) Sample a new association  $c_k^{(i)}$  with the following probabilities:
- Draw  $c_k^{(i)} = 0$  with probability  $\pi_0^{(i)}$
  - Draw  $c_k^{(i)} = 1$  with probability  $\pi_1^{(i)}$ .
  - Draw  $c_k^{(i)} = 2$  with probability  $\pi_2^{(i)}$ .
  - ...
  - Draw  $c_k^{(i)} = T$  with probability  $\pi_T^{(i)}$ .

Now it is easy to evaluate also the probabilities of data association hypotheses in the optimal importance distribution, because the probabilities are given by the terms  $\pi_j^{(i)}, j = 1, \dots, T$  conditional on each particle.

## 2.6 Algorithm Implementation

As shown in Section 2.2 the RBMCDA algorithm is a special case of the generic RBPF algorithm. However, due to the conditional independences between the targets, the full Kalman filter prediction and update steps for all targets can be reduced to single target predictions and updates. Actually, because the targets are a priori independent, conditional on the data associations  $c_k$  the targets will remain independent during tracking. This leads to the following simplifications to the RBPF computations:

- (1) The Kalman filter prediction steps can be done for each target in each particle separately. That is, we do not need to do Kalman filter prediction to the joint mean and covariance of all targets, but only to each target separately.
- (2) We can always use the optimal importance distribution, which was described in the previous section, as the importance distribution for the latent variables.
- (3) The marginalized measurement likelihoods can be computed for each target separately. Note that these likelihoods have already been computed for each target during the evaluation of the optimal importance distribution.
- (4) The measurement updates can be also performed for each target separately. This means that the Kalman filter update is actually performed only to one target in each particle.

## 3 RBMCDA with an Unknown Number of Targets

In the next sections we extend the RBMCDA algorithm to an unknown and time varying number of targets. The probabilistic construction of the model for an unknown number of targets, that is, the probability model for varying

state space dimension follows roughly the approach described in [9], which extends MHT to tracking an unknown number of targets. The idea is to assume that there is always a (very large) *constant number of targets*  $T_\infty$ . But *an unknown, varying number of them are visible (or alive)*, and they are the ones we are tracking. The visibility of targets is represented with a vector of discrete indicator variables.

The number  $T_\infty$  ensures that the joint distribution of the target states is a Gaussian distribution with constant dimensionality and thus a well defined probability distribution. The model is formulated such that we do not explicitly need to know the actual number of targets  $T_\infty$  as long as it is theoretically large enough.

As an extension to the MHT method, we construct a probabilistic stochastic process model for the births and deaths of the targets and show how Rao-Blackwellized particle filtering can be applied to this problem.

### 3.1 Filtering Model

- (1) *Clutter or false alarm measurements* have the same kind of model as in the RBMCDA model in Section 2.1.
- (2) *Target measurements* are linear Gaussian or non-linear EKF/UKF based as in the RBMCDA model in Section 2.1.
- (3) *Target dynamics* are linear Gaussian or EKF/UKF/IMM based as in the RBMCDA model in Section 2.1.
- (4) The target ( $c_k = j$ ) and clutter ( $c_k = 0$ ) association priors, in the case that births and deaths do not occur at the current time step, are known and can be modeled as an  $m$ th order Markov chain  $p(c_k | c_{k-m:k-1}, T_{k-m:k-1})$ , where  $T_{k-m:k-1}$  contains the number of targets at time steps  $k-m, \dots, k-1$ . For example, we might have a uniform prior over targets and clutter:

$$p(c_k | c_{k-m:k-1}, T_{k-m:k-1}) = \frac{1}{1 + T_{k-1}}. \quad (26)$$

- (5) Target births may happen only when a measurement is obtained, and in that case a birth happens with probability  $p_b$ . For simplicity, the model is defined such that a birth may happen only jointly with an association event, so that if there is no association to a newborn target, there is no birth.

This is equivalent to stating that the target state prior remains constant until the first measurement is associated, that is, the dynamic model does not affect the target state before the first measurement has been associated to the target. This indicates that it is sufficient to consider the time of the first associated measurement as the actual birth moment.

- (6) After associating a measurement with a target, the life time  $t_d$  (or time to death) of the target has probability density

$$t_d \sim p(t_d), \quad (27)$$

which can be, for example, an exponential or gamma distribution.

- (7) At the time of birth each target has a known Gaussian prior distribution (which can be non-informative)

$$p(\mathbf{x}_{k_0,j}) = N(\mathbf{x}_{0,j} \mid \mathbf{m}_{0,j}, \mathbf{P}_{k_0,j}). \quad (28)$$

### 3.2 Probabilities of Birth and Death

If a birth has occurred, it is assumed to be certain that the current measurement is associated to the newborn target:

$$p(c_k \mid \text{birth}) = \begin{cases} 1, & \text{if } c_k = T_{k-1} + 1 \\ 0, & \text{otherwise.} \end{cases} \quad (29)$$

In the case of no birth, the Markov model for data the associations applies:

$$p(c_k \mid \text{no birth}) = p(c_k \mid c_{k-m:k-1}). \quad (30)$$

The data association and birth events can be divided into the following cases with different probabilities:

- (1) A target is born and the measurement is associated with the newborn target:

$$\begin{aligned} b_k &= \text{birth} \\ c_k &= T_{k-1} + 1. \end{aligned} \quad (31)$$

- (2) A target is not born and the measurement is associated with one of the existing targets or with clutter:

$$\begin{aligned} b_k &= \text{no birth} \\ c_k &= j, \quad j = 0, \dots, T_k. \end{aligned} \quad (32)$$

- (3) Other events have zero probability.

Thus, given the associations  $c_{k-m:k-1}$  on the  $m$  previous steps, the joint distribution of the event  $b_k \in \{\text{no birth, birth}\}$  and the association  $c_k$  is given

as

$$p(b_k, c_k \mid c_{k-m:k-1}) = \begin{cases} p_b & , \text{ in case (1)} \\ (1 - p_b) p(c_k \mid c_{k-m:k-1}) & , \text{ in case (2)} \\ 0 & , \text{ in case (3)}, \end{cases} \quad (33)$$

where  $p_b$  (see Section 3.1) is the prior probability of birth.

The restriction of one data association per target on single time instance (see Section 2.3) in the case of unknown number of targets can be handled in the same manner as in the case of known number of targets. We simply assume that there is positive probability of detecting a newborn target on each step. That is, in addition to existing targets we model the possibility detection of a new target which has the detection probability  $p_b$ . This probability of detecting a new target is equivalent to the probability of birth, because we have defined the birth to be the event of detecting the target for the first time. However, the difference to the restriction of one association per target is that births may occur as many times as there are measurements on scan, not only once per scan.

The filtering model represented in Section 3.1, states that after associating a measurement with a target, the life time  $t_d$  of the target has the known probability density (27). Thus if the last association with target  $j$  was at time  $\tau_{k,j}$ , and on the previous time step  $t_{k-1}$  we sampled a hypothesis that the target is alive, then the probability that the target is dead at current time step  $t_k$  is

$$\begin{aligned} p(\text{death of } j \mid t_k, t_{k-1}, \tau_{k,j}) \\ = P(t_d \in [t_{k-1} - \tau_{k,j}, t_k - \tau_{k,j}] \mid t_d \geq t_{k-1} - \tau_{k,j}). \end{aligned} \quad (34)$$

### 3.3 Relationship to RBPF

The RBMCDA algorithm with an unknown number of targets fits to the RBPF framework, if the latent variable  $\boldsymbol{\lambda}_k$  contains the visibility indicator  $\mathbf{e}_k$  and the data association indicator  $c_k$  at the current time step

$$\boldsymbol{\lambda}_k = \{\mathbf{e}_k, c_k\}. \quad (35)$$

The visibility indicators and the data associations implicitly define the number of (visible) targets  $T_k$  at each time step.

Given that the targets are a priori unordered, there is a high permutation sym-



metry in the posterior distributions of the target states, visibility indicators and data association indicators. We can change the indices of any two targets, including the visibility indicators and data associations, and the probability of the configuration will remain the same. For this reason, we shall select one of the permutations arbitrarily and use it for representing all the permutations. This permutation is based on the times of the first associations with the targets. This does not change the model, because this is not a priori ordering, but is merely a way of selecting a compact representation for a very high number of redundant permutations.

- (1) The joint state  $\mathbf{x}_k$  contains the states of the  $T_\infty$  targets

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_{k,1} \\ \dots \\ \mathbf{x}_{k,T_\infty} \end{pmatrix}. \quad (36)$$

- (2) At the initial time step the targets have Gaussian prior distributions  $N(\mathbf{x}_{k,j} \mid \mathbf{m}_0, \mathbf{P}_0)$ . The model is constructed such that the invisible targets at any time step  $k$  (indicated by  $\mathbf{e}_k$ ) do not have a dynamic model. This means that the targets which have not yet become visible (have not been born yet) at any time step  $k$  have independent Gaussian prior distributions  $N(\mathbf{x}_{k,j} \mid \mathbf{m}_0, \mathbf{P}_0)$ . If we denote the sets of not visible and visible target indices with  $\mathcal{J}_0$  and  $\mathcal{J}_1$ , respectively, the joint prior distribution of all targets is of the form

$$\begin{aligned} p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) &= \prod_{j \in \mathcal{J}_1} N(\mathbf{x}_{k,j} \mid \mathbf{m}_{k,j}, \mathbf{P}_{k,j}) \\ &\times \prod_{j' \in \mathcal{J}_0} N(\mathbf{x}_{k,j'} \mid \mathbf{m}_0, \mathbf{P}_0). \end{aligned} \quad (37)$$

That is, the distribution of the visible targets is completely independent of the distributions of the invisible targets and thus it suffices to store only the states of the visible targets instead of all  $T_\infty$  targets. The joint distribution of all targets is still always theoretically a Gaussian distribution of dimension  $T_\infty$ .

- (3) When a target birth occurs, that is, a new target becomes visible (i.e., produces the first measurement) a new item in the indicator vector  $\mathbf{e}_k$  is set and the corresponding target prior distribution is updated (initialized) by the measurement. Because we only need to store one possible permutation from a high number of equivalent target permutations, we can add the new target to the first empty place in the indicator vector  $\mathbf{e}_k$ .
- (4) When a target dies, that is, becomes invisible again, the target distribution again becomes the prior and the target state is moved to the end of the joint state vector  $\mathbf{x}_k$  and indicator vector  $\mathbf{e}_k$ . The targets in the

vectors can be shifted such that the visible targets always remain in the beginning of the vectors.

- (5) The target dynamics, target measurements and clutter measurements are modeled in the same way as in the RBMCDA model of Section 2.1 and thus they fit into the RBPF framework easily.
- (6) By constructing a prior model for to births and deaths, we get the distribution

$$p(\mathbf{e}_k \mid \mathbf{e}_{k-1}), \quad (38)$$

which defines the dynamics of births and deaths. The data association model is of the form

$$p(c_k \mid c_{k-m:k-1}, \mathbf{e}_k), \quad (39)$$

and thus these two models together give a joint Markov chain model for the indicators:

$$\begin{aligned} p(\mathbf{e}_k, c_k \mid c_{k-m:k-1}, \mathbf{e}_{k-m:k-1}) \\ = p(c_k \mid c_{k-m:k-1}, \mathbf{e}_k) p(\mathbf{e}_k \mid \mathbf{e}_{k-1}), \end{aligned} \quad (40)$$

which is the form required by the RBPF model.

### 3.4 Evaluating and Sampling from the Optimal Importance Distribution

The possible events between two measurements  $\mathbf{y}_{k-1}$  and  $\mathbf{y}_k$  and at the association of measurement  $\mathbf{y}_k$  are:

- (1) Targets may die (indicated by elements of  $\mathbf{e}_k$ ):
  - (a) none of the targets dies
  - (b) one or more targets die
- (2)  $\mathbf{y}_k$  is associated with (indicated by  $c_k$ ):
  - (a) clutter
  - (b) one of the existing targets
  - (c) a newborn target

Death events are independent of the measurements. However, the two event families are related such that a new measurement  $\mathbf{y}_k$  can be associated only to the targets that have not died between the measurements  $\mathbf{y}_{k-1}$  and  $\mathbf{y}_k$ .

The model (40) assigns unique prior probabilities to each of the finite number of different events, but the problem is that the number of possible events grows exponentially with the number of targets. The combinatorial problem in the number of events is solely due to the exponential number of possible combinations of target deaths. There is no combinatorial problem in target births, because we can always use the sequential update scheme as discussed in Section 2.3. However, the purpose of the death model is only to remove the targets with which no measurements have been associated for a long time.

Because the death model is built only for serving this purpose without any physical meaning, it is not desirable to spend most of the computing power on it.

For the above reason we shall use an approximation where we restrict the possible number of deaths at each time step to one. This means that at each time step (measurement time) either none or exactly one of the targets dies. The probability of two or more death events between measurements is assumed to be zero. Note that the varying time step size is accounted already in the death model (34) and thus the survival rate will change only a bit, because we assume that there cannot be more deaths than there are measurements. We could overcome this restriction by performing redundant measurement update steps, for example, by generating artificial clutter measurements, which have prior probability one of being clutter.

Because due to restricting the number of deaths on each time step the number of events grows only linearly with the number of targets, we can construct an approximate prior distribution of births, deaths and associations as follows:

- (1) Enumerate all possible combinations of joint birth, (zero or one) deaths, and association events and compute probabilities for each of the combinations.
- (2) Normalize the list of events such that their probabilities sum to one.

For each combination of birth, death and association events there is a transition pair  $(\mathbf{e}_{1:k-1}, c_{1:k-1}) \rightarrow (\mathbf{e}_k, c_k)$  with a probability given by the above procedure. That is, we have an approximate representation of the distribution  $p(\mathbf{e}_k, c_k \mid \mathbf{e}_{1:k-1}, c_{1:k-1})$ .

The likelihood term  $p(\mathbf{y}_k \mid \mathbf{e}_k, c_k)$  can be computed similarly as in the case of a known number of targets (see Equation (20)). By multiplying each of the birth, death, and association combinations with the measurement likelihood and normalizing, we can form the optimal importance distribution similarly as in Section 2.5.

### 3.5 Data Representation

The algorithm state consists of a set of  $N$  particles, where each particle  $i$  at time step  $k$  contains the following:

$$\{c_{k-m+1:k}^{(i)}, \mathbf{e}_k^{(i)}, \mathbf{m}_{k,1}^{(i)}, \dots, \mathbf{m}_{k,j}^{(i)}, \dots, \mathbf{m}_{k,T}^{(i)}, \mathbf{P}_{k,1}^{(i)}, \dots, \mathbf{P}_{k,j}^{(i)}, \dots, \mathbf{P}_{k,T}^{(i)}, w_k^{(i)}\}, \quad (41)$$

where

- $c_{k-m+1:k}^{(i)}$  are the data association indicators of the time steps  $k-m+1, \dots, k$ .
- $\mathbf{e}_k^{(i)}$  is the life-indicator, which is a binary vector of length  $T_\infty$  indicating which of the targets are alive at current time step.
- $\mathbf{m}_{k,j}^{(i)}, \mathbf{P}_{k,j}^{(i)}$  are the mean and covariance of the target  $j$ , and they are conditional on the data association history  $c_{1:k}^{(i)}$ .
- $w_k^{(i)}$  is the importance weight of the particle.

The following information is also implicitly or explicitly stored for each particle:

$$\{T_k^{(i)}, \tau_{k,j}^{(i)}, \text{id}_{k,j}^{(i)}, \}, \quad (42)$$

where

- $T_k^{(i)}$  is the number of targets.
- $\tau_{k,j}^{(i)}$  is the time of the last measurement associated with target  $j$ .
- $\text{id}_{k,j}^{(i)}$  is a unique integer valued identifier, unique over all targets in all particles, which is assigned at the birth of the target.

### 3.6 Algorithm Implementation

Similarly to the case of RBMCDA (see, Section 2.6), because the targets are a priori independent, conditional on data associations  $c_k$  and indicators  $\mathbf{e}_k$ , the targets will also remain independent during tracking. This means that exactly the same simplifications to RBPF apply to the case of an unknown number of targets as to a known number of targets.

## 4 Simulations

### 4.1 Bearings Only Tracking of a Known Number of Targets

First we shall consider a classical bearings only multiple target tracking problem, which frequently arises in the context of passive sensor tracking. The simulation scenario is similar to that was presented in [1], but now the simulation includes clutter measurements and the restriction of one data association per target on single time instance is also modeled.

The dynamics of target  $j$  with the state vector  $\mathbf{x}_{j,k} = (x_{j,k} \ y_{j,k} \ \dot{x}_{j,k} \ \dot{y}_{j,k})^T$  can

be modeled with a discretized Wiener velocity model [6]

$$\begin{pmatrix} x_{j,k} \\ y_{j,k} \\ \dot{x}_{j,k} \\ \dot{y}_{j,k} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{j,k-1} \\ y_{j,k-1} \\ \dot{x}_{j,k-1} \\ \dot{y}_{j,k-1} \end{pmatrix} + \mathbf{q}_{k-1}, \quad (43)$$

where  $\mathbf{q}_{k-1}$  is the Gaussian process noise with moments

$$\begin{aligned} \mathbb{E}[\mathbf{q}_{k-1}] &= \mathbf{0} \\ \mathbb{E}[\mathbf{q}_{k-1}\mathbf{q}_{k-1}^T] &= \begin{pmatrix} \frac{1}{3}\Delta t^3 & 0 & \frac{1}{2}\Delta t^2 & 0 \\ 0 & \frac{1}{3}\Delta t^3 & 0 & \frac{1}{2}\Delta t^2 \\ \frac{1}{2}\Delta t^2 & 0 & \Delta t & 0 \\ 0 & \frac{1}{2}\Delta t^2 & 0 & \Delta t \end{pmatrix} q, \end{aligned}$$

where  $q$  is the spectral density of the noise. In the simulation, the value  $q = 0.1$  was used for both targets. The noise in an angular measurement from target  $j$  by sensor  $i$  can be modeled as Gaussian

$$\hat{\theta}_k = \arctan\left(\frac{y_{j,k} - s_y^i}{x_{j,k} - s_x^i}\right) + r_k, \quad (44)$$

where  $(s_x^i, s_y^i)$  is the position of sensor  $i$  and  $r_k \sim N(0, \sigma^2)$  with  $\sigma = 0.02$  radians.

Because the measurement model is non-linear we replace the Kalman filter in the data association algorithm with EKF. The uncertainty in data associations can be modeled by defining a variable  $c_k$ , which has the value  $c_k = j$  if the measurement at time step  $k$  is associated with target  $j$ .

The target detection probability is set to  $p_d = 80\%$  and the number of clutter measurements (uniform on range  $[-\pi, \pi]$ ) is Poisson-distributed with mean 5. The measurement data obtained from simulated sensors is shown in Figure 1. The initial distribution was on purpose selected such that all the four crossings of measurements from the two sensors contain some probability mass, and the distributions of the targets are two-modal as shown in Figure 2. The particles in the figure are a random sample drawn from the posterior distribution estimate, used for visualizing the distribution. The actual posterior distribution estimate is a mixture of Gaussians which is hard to visualize directly. The number of Monte Carlo samples used in the estimation method was 100.

Figure 3 shows the final tracking result, and it can be seen that in the beginning of the trajectory the posterior distribution is multimodal. Also the

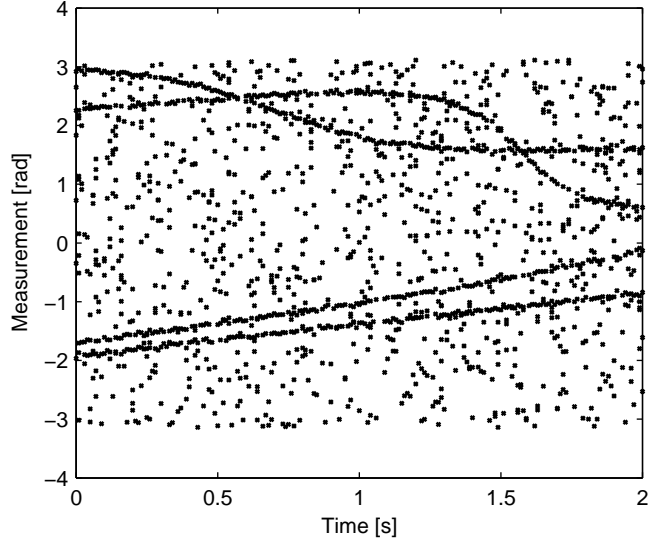


Fig. 1. Measurement data obtained from angular sensors.

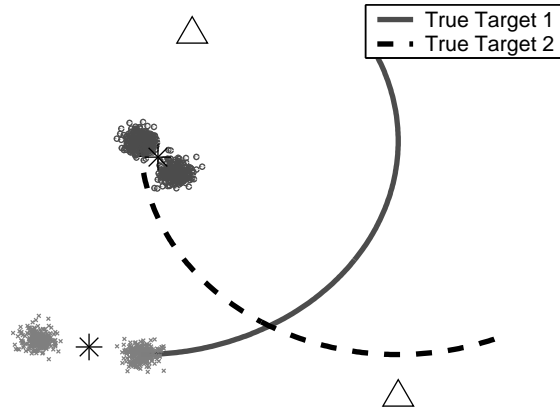


Fig. 2. The prior distributions of the targets. Half of the prior probability mass is located in the wrong sensor measurement crossings.

posterior distributions of the trajectories are a bit wider in the area between the sensors, because in that area the position uncertainty is higher in the direction of the line connecting the two angular sensors. Again, particles are used for visualizing the distribution, although the true posterior distribution estimate is a mixture of Gaussians.

Figure 4 shows the *smoothed* tracking result, which is an estimate where the distributions of all time steps are conditioned on all the measurements. This kind of an estimate can be easily calculated with (fixed interval) Kalman smoothers [2,6] and particle smoothers [4] also in the Rao-Blackwellized particle filtering case. Conditional on all the measurements the trajectory no longer contains multimodalities. Also the position uncertainty on the line connecting the sensors is lower.

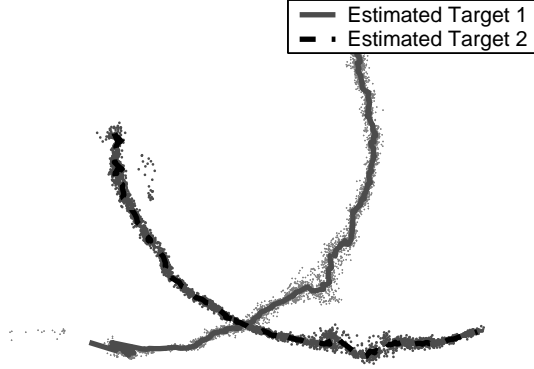


Fig. 3. Filter estimates for each time step. In the beginning of the trajectory the posterior distribution is multimodal. The multimodality can be seen from the two extra clouds of particles, which are located quite far away from the actual target trajectories. The higher position uncertainty on the line connecting the two sensors can be also seen.

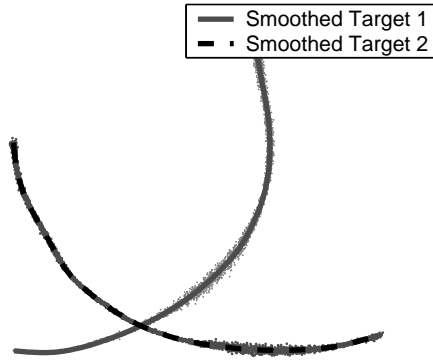


Fig. 4. Smoothed estimates do not have the multimodality in the beginning of the trajectory, because later measurements have resolved it. The position uncertainty on the line connecting the two sensors is also lower than in the filter estimates.

#### 4.2 Unknown Number of 1D Signals

In this section the algorithm extension to an unknown number of signals is simulated with the appearing and disappearing signals as described in Table 2. The true signals and simulated data are shown in the Figure 5. All the signals are modeled with discretized white noise acceleration models [6]

$$\begin{pmatrix} x_k \\ \dot{x}_k \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{k-1} \\ \dot{x}_{k-1} \end{pmatrix} + \mathbf{q}_{k-1}, \quad (45)$$

where  $x_k = x(t_k)$ ,  $\dot{x}_k = \dot{x}(t_k)$ , the sampling period is  $\Delta t = 1/100$ , and the process noise spectral density is  $q = 1/10$ . The signal conditioned measurements are modeled (and simulated) as measurements of the signal plus a white

<i>Signal</i>	<i>Appears</i>	<i>Disappears</i>
$x^{(1)}(t)$	$t = 0$	$t = 8$
$x^{(2)}(t)$	$t = 0$	$t = 15$
$x^{(3)}(t)$	$t = 1$	$t = 4$
$x^{(4)}(t)$	$t = 2$	$t = 5$
$x^{(5)}(t)$	$t = 5.5$	$t = 10$
$x^{(6)}(t)$	$t = 6$	$t = 15$

Table 2

Signal appearance and disappearance schedule in the 1D scenario with an unknown number of signals.

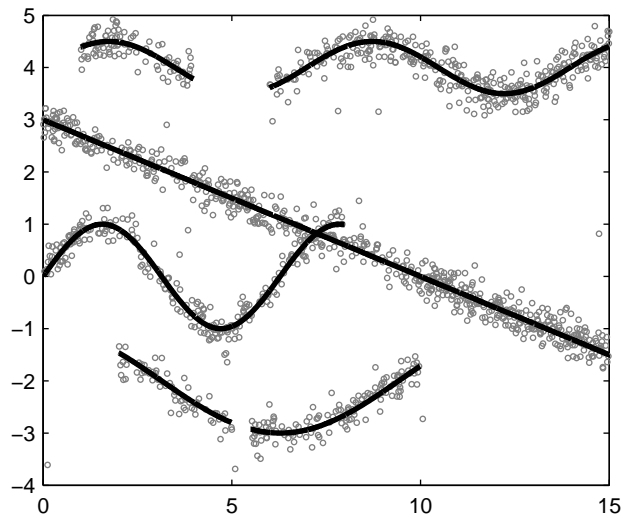


Fig. 5. Simulated data of the 1D scenario with an unknown number of signals.

Gaussian noise component

$$y_{k,j} = x^{(j)}(t_k) + r_k, \quad (46)$$

where  $r_k \in \mathcal{N}(0, 1/5^2)$ , given that the measurement is from signal  $j$ . Every measurement has an equal chance of originating from each of the visible signals and 1% chance of being a corrupted measurement uniformly distributed on the area  $[-5, 5]$ . The number of signals is unknown and the following model for the births and deaths is used:

- The prior probability of birth  $p_b = 1/100$ .
- A priori time to death  $t_d$  from the last data association has the gamma distribution

$$t_d \sim \text{Gamma}(t_d \mid \alpha, \beta), \quad (47)$$

with constant parameters  $\alpha$  and  $\beta$ .



The number of Monte Carlo samples used in estimation was  $N = 10$ . Figure 6 shows the result of filtering with parameters  $\alpha = 2$  and  $\beta = 1$ . The plotted result is the hypothesis contained in the particle with largest weight. As can be seen from Figure 7 there is slight delay after the disappearance of the signals before they disappear from estimation. The longest delay is in the signal that ends very near the other signal. Also the 1 time step gap in the lowest signal is not detected. Figure 8 shows the result of applying a Kalman smoother to the filtering result in the particle with the largest weight. This corresponds to the maximum a posteriori signal estimate. It can be seen that the estimation result follows the actual signal paths quite well except for the slight delays in signal disappearance.

Figures 9, 10 and 11 show the results when the disappearance model parameters were set to  $\alpha = 2$  and  $\beta = 1/10$ , which means that the signals disappear almost ten times faster than in the previous figures. It can be seen that in this case the signals are estimated to disappear and reappear also when there are random gaps in the signals due to uneven measurement times. This result is quite natural, because the model states that even quite small gaps in the measurement sequence (or actually in the signal) should be interpreted as disappearances and reappearances of the signal.

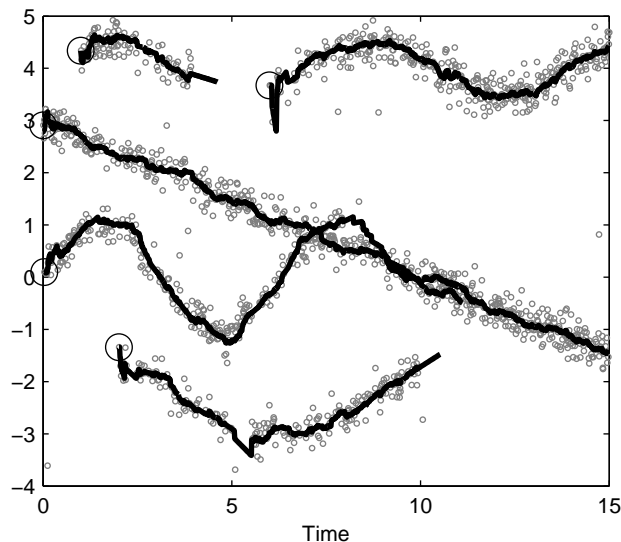


Fig. 6. Filtering result of the 1D scenario with an unknown number of signals and parameters  $\alpha = 2$  and  $\beta = 1$ . The circles represent the estimated starting points of the signals.

#### 4.3 Tracking an Unknown Number of Targets in 2D

In this section we demonstrate the algorithm in case of an unknown number of targets moving in 2D space. The prior model for the data associations is

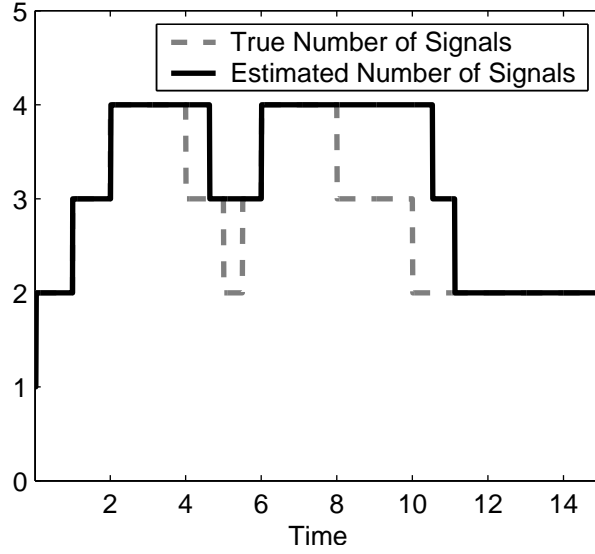


Fig. 7. Estimated number of signals in the 1D scenario with an unknown number of signals and parameters  $\alpha = 2$  and  $\beta = 1$ .

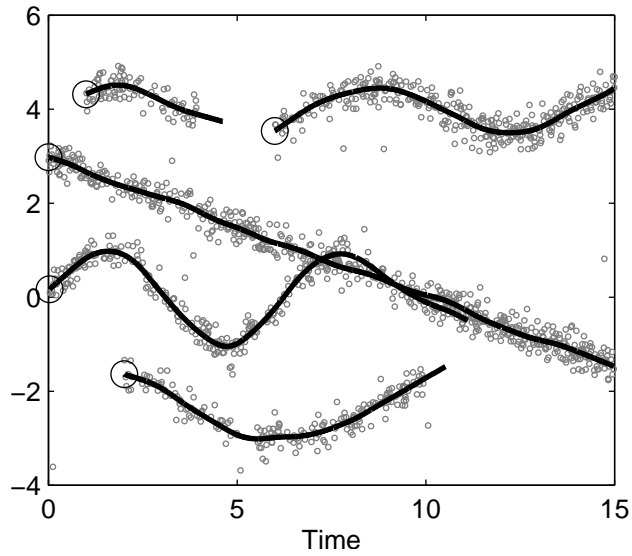


Fig. 8. Smoothing result of the 1D scenario with an unknown number of signals and parameters  $\alpha = 2$  and  $\beta = 1$ . The circles represent the estimated starting points of the signals.

defined such that only zero or one associations with each target on single scan is allowed.

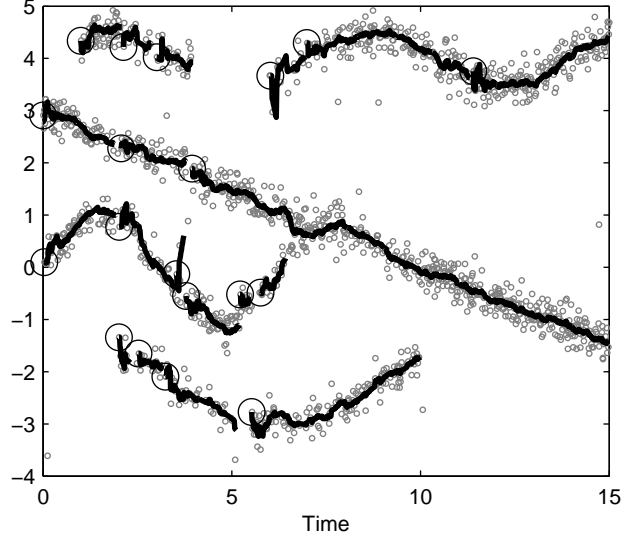


Fig. 9. Filtering result of the 1D scenario with an unknown number of signals and parameters  $\alpha = 2$  and  $\beta = 1/10$ . The circles represent the estimated starting points of the signals.

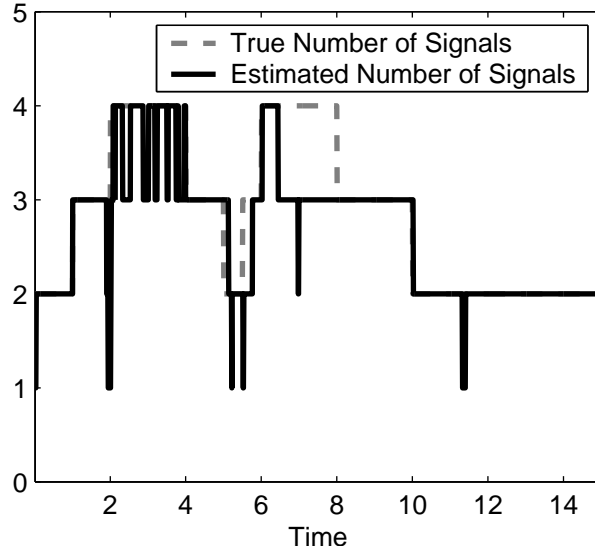


Fig. 10. Estimated number of signals in the 1D scenario with an unknown number of signals and parameters  $\alpha = 2$  and  $\beta = 1/10$ .

The dynamic model for the targets is

$$\begin{pmatrix} x_{j,k} \\ y_{j,k} \\ \dot{x}_{j,k} \\ \dot{y}_{j,k} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{j,k-1} \\ y_{j,k-1} \\ \dot{x}_{j,k-1} \\ \dot{y}_{j,k-1} \end{pmatrix} + \mathbf{q}_{k-1}, \quad (48)$$

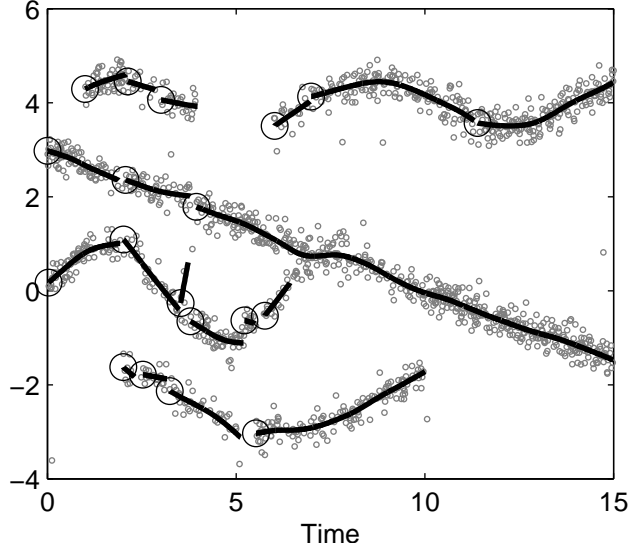


Fig. 11. Smoothing result of the 1D scenario with an unknown number of signals and parameters  $\alpha = 2$  and  $\beta = 1/10$ . The circles represent the estimated starting points of the signals.

where  $\mathbf{q}_{k-1}$  is process noise. The model for the measurements  $\mathbf{z}_k$  is

$$\begin{aligned} z_{1,k} &= x_k + r_{x,k} \\ z_{2,k} &= y_k + r_{y,k} \end{aligned} \quad (49)$$

where  $r_{x,k}, r_{y,k} \sim N(0, \sigma^2)$ . The sampling period  $\Delta t = 1/100$ , process noise  $q = 1/10$  in  $x$  and  $y$  directions, and measurement variance  $\sigma^2 = 1/20^2$ .

The detection probability of each target  $p_d = 95\%$  and at each time step a random number of clutter measurements on area  $[-2, 2] \times [-2, 2]$  is drawn from a Poisson distribution with mean 1. The clutter measurement prior is chosen to restrict maximum the number of data associations per target to one. The birth/death model parameters are set to  $p_b = 1/100$ ,  $\alpha = 2$  and  $\beta = 1/2$ . The number of Monte Carlo samples was  $N = 100$ .

Figures 12, 13 and 14 show the filtering results, estimated number of targets and smoothed results, respectively. Again, a slight delay in the estimated target disappearance can be seen after the actual disappearance, but yet the estimated trajectories follow the true trajectories quite well. Also the estimated number of targets matches the actual number of targets well.

The restriction of maximum number of data associations per target is very important for tracking performance, especially when the amount of clutter is significant. As we have seen 100 particles is enough for in this particular scenario, when the restriction is used. A quick test shows that if the restriction is not used, even with 1000 particles the result is much worse.

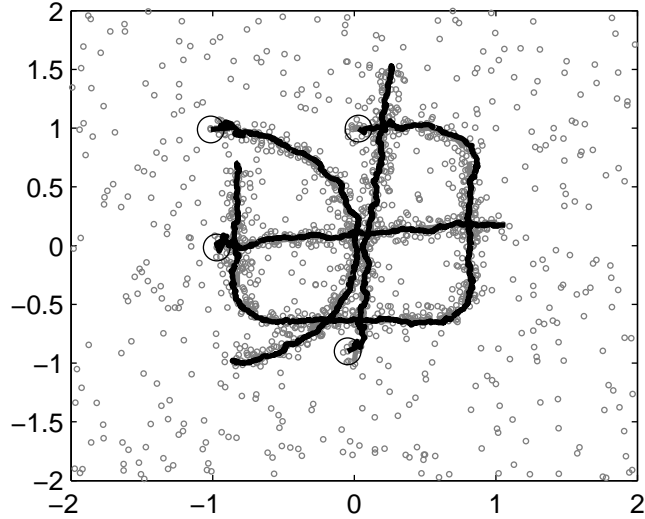


Fig. 12. Filtering result of the 2D scenario with an unknown number of targets. The circles represent the estimated starting points of the signals.

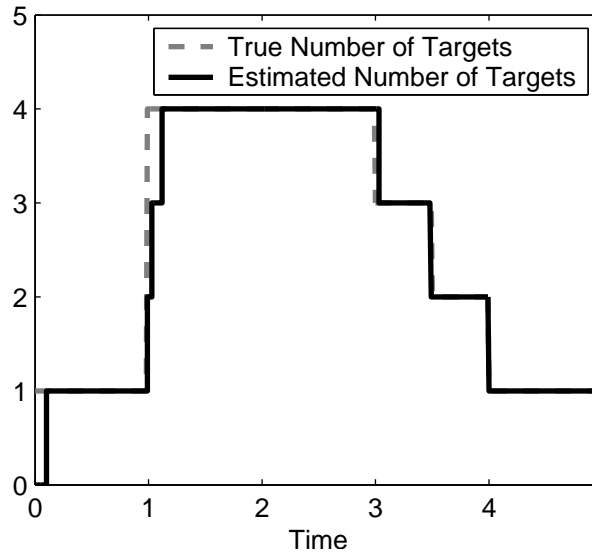


Fig. 13. Estimated number of targets in the 2D scenario with an unknown number of targets.

## 5 Conclusions and Discussion

In this article we have presented an extension of the Rao-Blackwellized Monte Carlo data association (RBMCD) algorithm [1] to tracking an unknown number of targets. The algorithm is based on constructing a Rao-Blackwellized particle filtering algorithm for the probabilistic model constructed for the target states, the data associations and the birth and death stochastic processes. Simulations were used for demonstrating the performance of the algorithm.

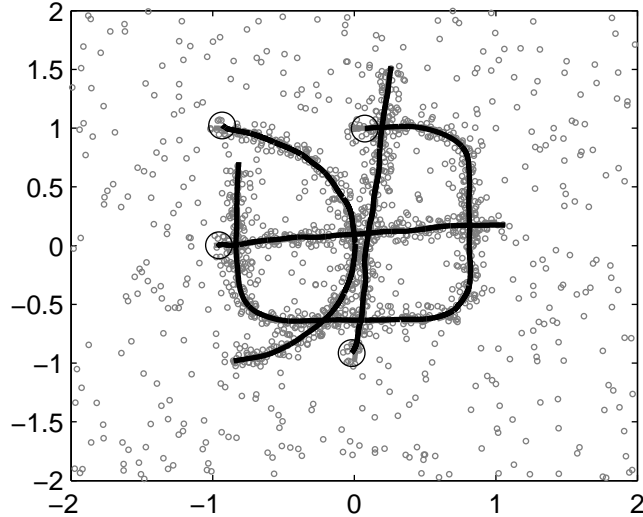


Fig. 14. Smoothing result of the 2D scenario with an unknown number of targets. The circles represent the estimated starting points of the signals.

The first extension to the RBMCDA in [1] is that the data associations are no longer required to be independent over target tracks. This allows considerable higher clutter densities, because the restriction of one data association per target on each time instance can be included into the model. As we have shown, this restriction can be modeled as an  $m$ th order Markov chain, which does not suffer from the combinatorial explosion of MHT. This is because finding the optimal (MAP) data association history is replaced with random sampling, which generates samples from the joint posterior distribution of target states and data associations and all possible data association on each time instance do not need to be evaluated. The price paid is that the minimum data association error (MAP) optimal solution is no longer available, but fortunately minimum mean squared error (MMSE) type estimates of target states can be still optimally computed from the posterior distribution.

The second extension is that the number of targets does not need to be known in advance and it can be time varying. Due to limiting the number of deaths between the measurements in the method, the computational complexity is of the same order as in the case of a known number of targets. This complexity should be approximately linear with respect to number of targets. Also, when the number of targets increases, the number of particles should be increased, which increases the computational complexity proportionally to the number of particles.

## Acknowledgments

The authors would like to thank Matti Vihola, Toni Tamminen and the anonymous reviewers for helpful comments.

## References

- [1] S. Särkkä, A. Vehtari, J. Lampinen, Rao-Blackwellized Monte Carlo data association for multiple target tracking, in: Proceedings of the Seventh International Conference on Information Fusion, Vol. I, 2004, pp. 583–590.
- [2] A. H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, 1970.
- [3] N. J. Gordon, D. J. Salmond, A. F. M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, in: IEEE Proceedings on Radar and Signal Processing, Vol. 140, 1993, pp. 107–113.
- [4] G. Kitagawa, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, Journal of Computational and Graphical Statistics 5 (1996) 1–25.
- [5] A. Doucet, N. de Freitas, N. Gordon (Eds.), Sequential Monte Carlo Methods in Practice, Springer, 2001.
- [6] Y. Bar-Shalom, X.-R. Li, T. Kirubarajan, Estimation with Applications to Tracking and Navigation, Wiley Interscience, 2001.
- [7] S. Blackman, R. Popoli, Design and Analysis of Modern Tracking Systems, Artech House Radar Library, 1999.
- [8] Y. Bar-Shalom, X.-R. Li, Multitarget-Multisensor Tracking: Principles and Techniques, YBS, 1995.
- [9] L. D. Stone, C. A. Barlow, T. L. Corwin, Bayesian Multiple Target Tracking, Artech House, Boston, London, 1999.
- [10] R. L. Streit, T. E. Luginbuhl, Maximum likelihood method for probabilistic multi-hypothesis tracking, in: Proceedings of Signal and Data Processing of Small targets, Vol. 2235, 1994, pp. 394–405.
- [11] N. Gordon, A hybrid bootstrap filter for target tracking in clutter, IEEE Transactions on Aerospace and Electronic Systems 33(1) (1997) 353–358.
- [12] R. Karlsson, F. Gustafsson, Monte Carlo data association for multiple target tracking, in: IEEE Target tracking: Algorithms and applications, The Netherlands, 2001.
- [13] C. Hue, J.-P. Le Cadre, P. Perez, The (MR)MTPF: particle filters to track multiple targets using multiple receivers, in: 4th International Conference on Information Fusion, Montreal, Canada, 2001.

- [14] R. P. Mahler, Random sets: Unification and computation for information fusion - a retrospective assessment, in: Proceedings of the Seventh International Conference on Information Fusion, Vol. I, 2004, pp. 1–20.
- [15] R. P. Mahler, Multitarget Bayes filtering via first-order multitarget moments, IEEE Transactions on Aerospace and Electronic Systems 39(4) (2003) 1152–1178.
- [16] B.-N. Vo, S. Singh, A. Doucet, Sequential Monte Carlo implementation of the PHD filter for multiple-target tracking, in: Proceedings of the International Conference on Information Fusion, 2003, pp. 792–799.
- [17] H. Sidenbladh, Multi-target particle filtering for the probability hypothesis density, in: Proceedings of the International Conference on Information Fusion, 2003, pp. 1110–1117.
- [18] C. Kreucher, K. Kastella, A. O. Hero III, Multitarget tracking using a particle filter representation of the joint multitarget density, in: Proceedings of Signal and Data Processing of Small targets, Vol. 5204, 2003.
- [19] A. Doucet, B.-N. Vo, C. Andrieu, M. Davy, Particle filter for multi-target tracing and sensor management, in: Proceedings of the Fifth International Conference on Information Fusion, 2002, pp. 474–481.
- [20] M. Isard, J. MacCormick, BraMBLe: A Bayesian multiple-blob tracker, in: Proceedings of International Conference on Computer Vision, Vol. 2, 2001, pp. 34–41.
- [21] D. J. Ballantyne, H. Y. Chan, M. A. Kouritzin, A branching particle-based nonlinear filter for multi-target tracking, in: Proceedings of the Fourth International Conference on Information Fusion, 2001.
- [22] N. de Freitas, C. Andrieu, Højen-Sørensen, M. Niranjan, A. Gee, Sequential Monte Carlo methods for neural networks, in: Sequential Monte Carlo Methods in Practice, 2001, pp. 359–379.
- [23] B. Ristic, S. Arulampalam, N. Gordon, Beyond the Kalman Filter, Artech House, 2004.
- [24] J. S. Liu, R. Chen, Blind deconvolution via sequential imputations, Journal of the American Statistical Association 90(430) (1995) 567–576.
- [25] H. Akashi, H. Kumamoto, Random sampling approach to state estimation in switching environments, Automatica 13 (1977) 429–434.
- [26] G. Casella, C. P. Robert, Rao-Blackwellisation of sampling schemes, Biometrika 83(1) (1996) 81–94.
- [27] P. Fearnhead, P. Clifford, On-line inference for hidden Markov models via particle filters, J. R. Statist. Soc. B 65(4) (2003) 887–899.
- [28] E. Punskeya, A. Doucet, W. J. Fitzgerald, On the use and misuse of particle filtering in digital communications, in: Proceedings of EUSIPCO, 2002.



- [29] S. J. Julier, J. K. Uhlmann, Unscented filtering and nonlinear estimation, Proceedings of the IEEE 92(3) (2004) 401–422.