Nationaal Lucht- en Ruimtevaartlaboratorium

National Aerospace Laboratory NLR

Executive summary

Exact Bayesian and particle filtering of stochastic hybrid systems

Problem area
In literature on Bayesian filtering of stochastic hybrid systems most studies are limited to Markov jump systems. The main exceptions are approximate Bayesian filters for semi-Markov jump linear systems. These studies showed that nonlinear filtering becomes much more challenging under non-Markov jumps. This challenge however does not apply to particle filtering of stochastic hybrid systems. In practice, non-Markov jumps rather are the rule, not the exception. For example, on an airport, the probability at which a taxiing aircraft makes a maneuver depends heavily on its position; e.g. when taxiing near a crossing on the airport, the probability of starting a turn is relatively high, whereas outside these areas this probability may be very small. A similar difference applies to the probability of an aircraft making a turn when it is flying near a waypoint versus flying halfway two waypoints. Similar non-Markov jump behavior also applies to other traffic modalities, and to any other intelligently controlled system. Nevertheless in target tracking, particle filtering studies have continued to focus on Markov jump systems

Description of work
This report considers filtering of stochastic hybrid systems that go beyond the well known Markov jump system. First the non-Markovian jump system studied in this report is formally defined. Next the exact Bayesian filter recursion for this system is developed and the implication of going beyond Markov jumps is shown. Subsequently a novel particle filter for jump nonlinear systems is developed which is referred to as the Interacting Multiple Model particle filter (IMMPF). For comparison Monte Carlo simulations are performed.

Results and conclusions
Through Monte Carlo simulations IMMPF has been tested and compared with standard Particle Filter (PF) and IMM. The results show that the IMMPF performs very well, even in cases where the standard PF or IMM have problems.

Applicability
The applicability of the work comprises the implementation of the resulting filtering algorithms in a multitarget tracker, in particular the Advanced suRveillance Tracker And Server ARTAS.

Report no.
NLR-TP-2006-691

Author(s)
H.A.P. Blom, E.A. Bloem

Report classification
Unclassified

Date
September 2006

Knowledge area(s)
Advanced (sensor-) information processing

Descriptor(s)
Stochastic hybrid systems
Bayesian filtering
Particle filtering
State dependent switching
Jump-nonlinear systems
Non-Markov jumps
Maneuvering target tracking

Nationaal Lucht- en Ruimtevaartlaboratorium, National Aerospace Laboratory NLR

Anthony Fokkerweg 2, 1059 CM Amsterdam,
P.O. Box 90502, 1006 BM Amsterdam, The Netherlands
Telephone +31 20 511 31 13, Fax +31 20 511 32 10, Web site: www.nlr.nl

Nationaal Lucht- en Ruimtevaartlaboratorium
National Aerospace Laboratory NLR

NLR-TP-2006-691

Exact Bayesian and particle filtering of stochastic hybrid systems

H.A.P. Blom and E.A. Bloem

This report contains an article to appear in IEEE Transaction on Aerospace and Electronic Systems.

This report may be cited on condition that full credit is given to NLR and the authors.

Customer:	National Aerospace Laboratory NLR
Working Plan number:	2005 AT.1.C
Owner:	National Aerospace Laboratory NLR
Division:	Air Transport
Distribution:	Unlimited
Classification title:	Unclassified
	September 2006

Approved by:

Author	Reviewer	Managing department
8-12-06	Anonymous peer reviewers	22/12/06

Exact Bayesian and Particle Filtering of Stochastic Hybrid Systems

Henk A.P. Blom Senior Member IEEE & Edwin A. Bloem

Abstract—**The standard way of applying particle filtering to stochastic hybrid systems is to make use of hybrid particles, where each particle consists of two components, one assuming Euclidean values, and the other assuming discrete mode values. This paper develops a novel particle filter for a discrete-time stochastic hybrid system. The novelty lies in the use of the exact Bayesian equations for the conditional mode probabilities given the observations. Therefore particles are needed for the Euclidean valued state component only. The novel particle filter is referred to as the Interacting Multiple Model (IMM) particle filter because it incorporates a filter step which is of the same form as the interaction step of the IMM algorithm. Through Monte Carlo simulations, it is shown that the IMM particle filter has significant advantage over the standard particle filter, in particular for situations where conditional switching rate or conditional mode probabilities have small values.**

Keywords: Stochastic hybrid systems, Bayesian filtering, Particle filtering, state dependent switching, jump-nonlinear systems, non-Markov jumps, maneuvering target tracking.

I. INTRODUCTION

The paradigm of particle filtering has stimulated a renewed interest in exact Bayesian filtering of non-linear stochastic systems. In line with this, [1] recently developed the exact Bayesian filter for Markov jump non-linear systems. The aim of the current paper is to extend the exact Bayesian filter characterization to the much larger class of non-Markov jump non-linear systems, and subsequently to exploit this exact Bayesian characterization for the development of a novel particle filter for stochastic hybrid systems.

In literature on Bayesian filtering of stochastic hybrid systems most studies are limited to Markov jump systems. The main exceptions are approximate Bayesian filters for semi-Markov jump linear systems [2] – [4]. These studies showed that nonlinear filtering becomes much more challenging under non-Markov jumps. This challenge however does not apply to particle filtering of stochastic hybrid systems [5]. Hence there is no reason to continue focusing on Markov jump systems. In practice, non-Markov jumps rather are the rule, not the exception. For example, on an airport, the probability at which a taxiing aircraft makes a maneuver depends heavily on its position; e.g. when taxiing near a crossing on the airport, the probability of starting a turn is relatively high, whereas outside these areas this probability may be very small. A similar difference applies to the probability of an aircraft making a turn when it is flying near a way-point versus flying halfway two way-points. Similar non-Markov jump behavior also applies to other traffic modalities, and to any other intelligently controlled system. Nevertheless in target tracking, particle filtering studies have continued to focus on Markov jump systems [6] – [19].

In nonlinear filtering studies, the Sampling Importance Resampling (SIR) based approach [20] has surfaced as the baseline particle filter. It has shown to form an elegant and general approach towards the numerical evaluation of the conditional density solution of the Chapman-Kolmogorov-Bayes (CKB) filter recursion [21] – [24]. For a large class of problems it has been shown that increasing the number of particles ensures weak sense convergence of the approximation density to the exact conditional density at a rate that is independent of or increases linearly with the state dimension of the process to be estimated [25] – [27]. This means that the SIR particle filter significantly relaxes the curse of dimensionality of CKB filtering. Practically, the SIR particle filter has found a large variety of useful applications, including some where established non-linear filtering approaches do not work at all, such as the visual tracking example of [28] and the track-before-detect example of [29]. Because of its generality and theoretical validity, the SIR particle filter has become a useful reference in numerically approximating CKB filter performance [18].

The SIR particle filter is also capable in approximating the CKB equations of a stochastic hybrid Markov process $\{x_t, \theta_t\}$, with x_t assuming values in $\mathbb{R}^n$, and θ_t assuming

Manuscript received October 25, 2004; revised October 7, 2005; released for publication March, 2006

This work was supported in part by the European Commission project HYBRIDGE, IST-2001-32460.
Both authors are with the National Aerospace Laboratory NLR, P.O. Box 90502, 1006BM Amsterdam, The Netherlands (phone: +31 20 5113544; fax: +31 20 5113210; e-mail: blom@nlr.nl, bloem@nlr.nl).

values in a finite set $\mathbb{M}$ of possible modes [5]. Let $\{x_t, \theta_t\}$ be the hidden state process to be estimated from noisy observations, then the SIR particle filter use N_P particles, particle *j* of which has two components (x_t^j, θ_t^j) at moment *t*, with x_t^j assuming an Euclidean value and θ_t^j assuming a mode value. Each SIR particle filter cycle from *t*-1 to *t* consists of three steps:

- Evolution: For each of the N_p particles at moment *t*-1, draw a new hybrid particle $(\bar{x}_t^j, \bar{\theta}_t^j)$ according to the Chapman-Kolmogorov transition kernel;
- Correction: For each of the N_p particles evaluate $\bar{\mu}_t^j$ as the likelihood of the measurement at moment *t*; given $(\bar{x}_t^j, \bar{\theta}_t^j)$ and normalize the resulting $\bar{\mu}_t^j$'s.
- Resampling: Draw N_p independently identically distributed (i.i.d.) hybrid particle values (x_t^j, θ_t^j), with from the sum of $\bar{\mu}_t^j$ weighted Dirac measures at $(\bar{x}_t^j, \bar{\theta}_t^j)$.

Successful application of this hybrid state version of the SIR particle filter has for example been shown for target tracking [6] – [7], signal processing [11] and failure monitoring and diagnosis [5]. In none of these example applications the conditional mode probabilities and/or switching rates assumed very small values. Otherwise, there might have been very few (or zero) particles for one or more of the mode values, and then the empirical density spanned by the particles with such a mode value does not form an accurate approximation of the mode probability or the corresponding mode conditional density. A brute force approach in compensating for sample degeneracy would be to increase the number of particles. The aim of this paper is to develop a more elegant approach in improving the SIR particle filter for stochastic hybrid systems.

Degeneration of samples is a well known phenomenon in particle filtering. Hence in literature various mitigation approaches have been developed, i.e.:

- Less frequent resampling, i.e. perform resampling only when the ”effective sample size” drops below a certain threshold [30].
- Better resampling methods. [23] showed that a resampling method which systematically draws one random sample per group of samples having weight of $1/N_p$ reduces the chance of sample degeneracy. Another good alternative is residual resampling [30]. A complementary advantage of systematic and residual resampling is that they work well without the need of ordering the samples according to their weights.
- Boosting diversity by regularization, i.e. prior to resampling, replace each Dirac measure in the empirical distribution by an absolutely continuous distribution. This approach has largely been developed for particle filtering of stochastic processes that, without regularization, do not satisfy the regularity conditions under which the weak convergence to the exact conditional density is satisfied [12]. A valuable alternative is to insert a Monte Carlo Markov Chain (MCMC) move in each SIR cycle [31], [11].
- Importance density based resampling. This means that particle resampling anticipates how the weights per particle will evolve over the next prediction and correction steps. Each new particle sample receives a weight μ_t^j that compensates for this importance resampling. The optimal importance density has been characterized by Doucet et al. [32]. Because exact implementation is hard, approximations are needed. Best known is the auxiliary resampling method by [33]. Other relevant approximate importance resampling methods have been developed by [32], [34] – [36].
- Rao-Blackwellization. This approach collects all state components into two vectors. One or more state components that satisfy a linear equation and that do not appear in the measurement equation, are placed in the first vector. The other state components are placed in the second vector. For the second vector, all filtering steps are performed through particles [37]. Per particle, the conditional density of the first vector follows from a Kalman filter, in which the value of the particle plays the role of the measurement, e.g. [9], [32], [38] – [40].

Typically an improvement of the SIR particle filter is used to reduce the number of particles needed to realize good filtering performance. Recently, [19] compared the effectiveness of importance density resampling and Rao-Blackwellization SIR improvement methods with the baseline SIR particle filter. For some specific examples of tracking a frequently maneuvering target in clutter, importance density resampling in combination with less frequent resampling showed to be most effective in mitigating sample degeneracy and without significantly increasing the computational load per particle.

All improvements of the SIR particle filter mentioned above have been developed with focus on Euclidean valued state estimation. This suggests there is room for improvement of the SIR particle filter with focus on the finite valued state component. To make this work, we apply a novel kind of Rao-Blackwellization to a stochastic hybrid process. We collect all Euclidean valued components in one vector, and place the mode component only in the other vector. Subsequently we perform particle filtering to the Euclidean valued vector and apply the exact filter equations to the mode probabilities. Exactly the same splitting of mode and state components is used by the Rao-Blackwellization application to a Markov jump linear system [32], but with the allocation of particles and exact filter equations the other way around. Independently of each other, [17] and [41] have recognized the value of this approach, with the latter including non-Markov jumps. Their particle filters for estimating the conditional densities of x_t given the mode θ_t are equivalent. However, the equations for the estimation of conditional mode probabilities differ as a result of the Markov/non-Markov jump difference. The current paper gives a full exposition of the more general approach developed by [41].

The paper is organized as follows. Section II formally defines the non-Markovian jump system studied in this paper. Section III develops the exact Bayesian filter recursion for this system, and shows the implication of going beyond Markov jumps. Section IV develops a novel particle filter for jump nonlinear systems. This novel particle filter has an interaction step which follows from the initial derivation [42] of the Interacting Multiple Model (IMM) filter algorithm. For this reason, the novel particle filter is referred to as the IMM particle filter. Section V performs Monte Carlo simulations to compare the novel particle filters with the SIR particle filter and with the IMM; for a fair comparison with IMM, the example considered is a Markov-jump linear system. Through parameter variation, the different effects of rare and non-rare switching probability is discussed. Finally, Section VI draws conclusions.

II. FORMULATION OF THE PROBLEM

We consider the following system of stochastic difference equations, on $[0,T]$, $T<\infty$,

$$x_t = a\left(\theta_t, x_{t-1}, w_t\right) \tag{1}$$

$$\theta_t = c\left(\theta_{t-1}, x_{t-1}, u_t\right) \tag{2}$$

$$y_t = h\left(\theta_t, x_t, v_t\right) \tag{3}$$

where the pair (θ_t, x_t) represents the hybrid system state, and y_t represents the observation at moment t, $\{w_t\}$ and $\{v_t\}$ are independent sequences of i.i.d. standard Gaussian variables of dimension n' and m' respectively, $\{u_t\}$ is an $\{w_t, v_t\}$-independent sequence of i.i.d. standard uniform random variables, $\{w_t, v_t, u_t\}$ is independent of the $\mathbb{M}\times\mathbb{R}^n$ valued initial condition (θ_0, x_0), with $\mathbb{M}$ a set of M discrete modes. Furthermore, a and h are measurable mappings of $\mathbb{M}\times\mathbb{R}^n\times\mathbb{R}^{n'}$ into $\mathbb{R}^n$ and $\mathbb{M}\times\mathbb{R}^n\times\mathbb{R}^{m'}$ into $\mathbb{R}^m$ respectively, and c is a measurable mapping of $\mathbb{M}\times\mathbb{R}^n\times[0,1]$ into $\mathbb{M}$. The mappings a, c and h are time-invariant for notational simplicity only.

The filtering problem is to estimate the joint conditional density-probability $p_{x_t,\theta_t|Y_t}(x,\theta)$, $x\in\mathbb{R}^n, \theta\in\mathbb{M}$, of the pair (x_t, θ_t) given the sequence of observations $Y_t = \{y_s; s\le t\}$. This problem is addressed in the sequel of this paper. In preparation for this, in this section we characterize some properties of the solution $\{\theta_t, x_t, y_t\}$ of system (1)-(3).

Firstly, by an iterative way of working from $t=0$ to $t=1$, then to $t=2$, and so on to $t=T$, it can be verified that for every initial condition (θ_0, x_0) the system of equations (1)-(3) has a unique measurable solution $\{\theta_t, x_t, y_t\}$ assuming values in $\mathbb{M}\times\mathbb{R}^n\times\mathbb{R}^m\times[0,T]$, $T<\infty$.

Secondly, according to equation (2), θ_t is defined as a function of θ_{t-1}, x_{t-1} and u_t. This implies that the evolution of the process $\{\theta_t\}$ depends of $\{x_t\}$. Hence the process $\{\theta_t\}$ is not a Markov process.

Thirdly, the joint process $\{x_t, \theta_t\}$ can be shown to be a Markov process. Substituting (2) into (1):

$$x_t = a\left(c\left(\theta_{t-1}, x_{t-1}, u_t\right), x_{t-1}, w_t\right) \tag{4}$$

Together with (2), this implies that (θ_t, x_t) is a measurable function of θ_{t-1}, x_{t-1}, u_t and w_t, for every $t>0$. By a repeated application of this functional relation, it follows that for every $s>t>0$, (θ_s, x_s) is a measurable function of θ_t, x_t, $(u_t,...,u_s)$ and $(w_t,...,w_s)$. This implies that the joint process $\{x_t, \theta_t\}$ satisfies the Markov property.

Fourthly, we characterize the transition probability $\Pi_{\eta\theta}$ of a jump from $\theta_{t-1}=\eta$ to $\theta_t=\theta$ in terms of properties of x_{t-1} and u_t, Application of the total probability theorem and subsequent evaluation, using (2), yields:

$$\begin{aligned} p_{\theta_t|\theta_{t-1},x_{t-1}}(\theta\mid\eta,x) &= \int_{\mathbb{R}} p_{u_t,\theta_t|\theta_{t-1},x_{t-1}}(u,\theta\mid\eta,x)du \\ &= \int_{\mathbb{R}} p_{\theta_t|\theta_{t-1},x_{t-1},u_t}(\theta\mid\eta,x,u)\,p_{u_t}(u)du \\ &= \int_{[0,1]} \chi\left(\theta, c(\eta,x,u)\right)p_{u_t}(u)du \end{aligned} \tag{5}$$

where χ a 0-1 indicator with $\chi(\theta,\theta')=1$ iff $\theta=\theta'$. This means that it makes sense to define the state-dependent mode transition probabilities $\Pi_{\eta\theta}(x)$, as follows:

$$\Pi_{\eta\theta}(x) \triangleq p_{\theta_t|\theta_{t-1},x_{t-1}}(\theta\,|\,\eta,x), \;\; (\theta,\eta)\in\mathbb{M}^2, \; x\in\mathbb{R}^n$$
$$= \int_{[0,1]} \chi(\theta, c(\eta,x,u))\, p_{u_t}(u)du \qquad (6)$$

Fifthly, from (6) we can see that if $c(\theta,x,u)$ is x-invariant then $\Pi_{\eta\theta}(x)$ is x-invariant and $\{\theta_t\}$ is a finite state Markov process.

In return for loosing $\{\theta_t\}$'s Markov property, the advantage of working with an x-variant $c(\theta,x,u)$ in (2) is the capability to define a mode valued jump process $\{\theta_t\}$ that incorporates stochastic hybrid processes where both the process $\{\theta_t\}$ influences the evolution of process $\{x_t\}$ and the process $\{x_t\}$ influences the evolution of process $\{\theta_t\}$. In order to illustrate this effect, we give two examples where mode switching depends on the evolution of $\{x_t\}$.

Example 1: We consider an open tank receiving an input flow of liquid from above. Liquid can be taken out / left in the tank by switching a pump on/off. The pumping mode is captured by a process $\{\theta_t\}$ assuming values in *{On,Off}*. The level of liquid in the tank is captured by an Euclidean valued process $\{x_t\}$. Assume the desired reference level of liquid in the tank is R. In order to avoid that the pump is rapidly switching on and off, the switching takes into account a hysteresis of strictly positive size H in level. Hence a pump in *On*-mode is switched to *Off* as soon as the level x_t, hits the level R-H. Similarly, an *Off*-mode pump is switched to *On* as soon as the level x_t hits the level R+H. Hence the mapping c satisfies:

$$\begin{aligned} c(\eta,x,u) &= Off, \text{ if } \eta = Off \text{ AND } x < R+H \\ &= On, \text{ if } \eta = Off \text{ AND } x \geq R+H \\ &= On, \text{ if } \eta = On \text{ AND } x > R-H \\ &= Off, \text{ if } \eta = On \text{ AND } x \leq R-H \end{aligned}$$

Substituting this into eq. (6) and evaluation yields:

$$\begin{aligned} \Pi_{\eta\theta}(x) &= 1, \;\text{ if } x \geq R+H, \; \eta = Off, \; \theta = On \\ &= 0, \;\text{ if } x < R+H, \; \eta = Off, \; \theta = On \\ &= 1, \;\text{ if } x \leq R\text{-}H, \; \eta = On, \; \theta = Off \\ &= 0, \;\text{ if } x \;> R\text{-}H, \; \eta = On, \; \theta = Off \end{aligned}$$

Example 2: For the same tank example we now assume pump switching errors happen at a rate ε. This can be modeled by the c-mapping as follows:

$$\begin{aligned} c(\eta,x,u) &= Off, \text{ if } \eta = Off \;, x < R+H, u > \varepsilon \\ &= On, \text{ if } \eta = Off \;, x < R+H, u \leq \varepsilon \\ &= On, \text{ if } \eta = Off \;, x \geq R+H, u > \varepsilon \\ &= Off, \text{ if } \eta = Off \;, x \geq R+H, u \leq \varepsilon \\ &= On, \text{ if } \eta = On \;, x > R-H, u > \varepsilon \\ &= Off, \text{ if } \eta = On \;, x > R-H, u \leq \varepsilon \\ &= Off, \text{ if } \eta = On \;, x \leq R-H, u > \varepsilon \\ &= On, \text{ if } \eta = On \;, x \leq R-H, u \leq \varepsilon \end{aligned}$$

For the transition probability this implies:

$$\begin{aligned} \Pi_{\eta\theta}(x) &= 1-\varepsilon, \;\text{ if } x \geq R+H, \; \eta = Off, \; \theta = On \\ &= \varepsilon, \qquad \text{if } x < R+H, \; \eta = Off, \; \theta = On \\ &= 1-\varepsilon, \;\text{ if } x \leq R\text{-}H, \; \eta = On, \; \theta = Off \\ &= \varepsilon, \qquad \text{if } x \;> R\text{-}H, \; \eta = On, \; \theta = Off \end{aligned}$$

In example 1 the mode switching is determined in a deterministic way by the evolution of $\{x_t\}$, and the i.i.d. sequence $\{u_t\}$ has no influence on this at all. This $\{x_t\}$-determined mode switching character is totally different from the well known Markov kind of $\{u_t\}$-determined mode switching. In example 2 the pure $\{x_t\}$-determined mode switching behavior is combined with some random $\{u_t\}$-determined mode switching.

In [5] it is demonstrated that filtering for stochastic hybrid systems, which incorporate the example 1 kind of x_t-dependent switching, is well handled by the SIR particle filter. For example 2, however, a SIR particle filter with N_p particles works well when $\varepsilon = 0$ or when $\varepsilon >> 1/N_p$, but not when $0 < \varepsilon < 1/N_p$. Such discontinuity in behavior is rather unnatural, and asks for the development of an appropriate improvement of the SIR particle filter. Prior to developing an adequate solution for this in section IV, we first characterize the exact Bayesian filter recursion from the general CKB equations.

III. EXACT BAYESIAN FILTER

In this section we develop the exact recursive equations for the joint conditional density-probability $p_{x_t,\theta_t|Y_t}(x,\theta)$, where Y_t denotes the σ-algebra generated by measurements y_t up to and including moment t. For short we refer to $p_{x_t,\theta_t|Y_t}(x,\theta)$ as the conditional density. Because $\{x_t,\theta_t\}$ is a Markov process, the characterization of this conditional density consists of two steps:

- Chapman-Kolmogorov (CK) evolution from $t-1$ to t, i.e. characterize $p_{x_t,\theta_t|Y_{t-1}}(x,\theta)$ as a function of $p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\theta)$.
- Bayes measurement update, i.e. characterize $p_{x_t,\theta_t|Y_t}(x,\theta)$ as a function of $p_{x_t,\theta_t|Y_{t-1}}(x,\theta)$.

The Bayes measurement update works similar as for measurements of a Markov jump non-linear system, e.g. [1]. For the characterisation of the CK step of the non-Markov jump system, however, this derivation path falls short. Hence, we follow other derivation paths. One derivation path uses the law of total probability, and is based on [41]. The second derivation path uses the Chapman-Kolmogorov equation directly, and is based on [14]. First we show the law of total probability derivation path, then we formally state the result in Theorem 1, including the assumptions adopted. Next we formally prove Theorem 1 following the Chapman-Kolmogorov derivation path.

By law of total probability

$$\begin{aligned} p_{x_{t-1},\theta_t|Y_{t-1}}(x,\theta) &= \sum_{\eta\in\mathbb{M}} p_{x_{t-1},\theta_t,\theta_{t-1}|Y_{t-1}}(x,\theta,\eta) = \\ &= \sum_{\eta\in\mathbb{M}} p_{\theta_t|x_{t-1},\theta_{t-1},Y_{t-1}}(\theta\,|\,x,\eta)\,p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\eta) \end{aligned} \tag{7}$$

Because u_t is independent of Y_{t-1}, from eq. (2) follows that θ_t is conditionally independent of Y_{t-1} given (θ_{t-1},x_{t-1}). Hence (7) yields

$$p_{x_{t-1},\theta_t|Y_{t-1}}(x,\theta) = \sum_{\eta\in\mathbb{M}} p_{\theta_t|x_{t-1},\theta_{t-1}}(\theta\,|\,x,\eta)\,p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\eta)$$

$$= \sum_{\eta\in\mathbb{M}} \Pi_{\eta\theta}(x)\,p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\eta) \tag{8}$$

Also by law of total probability:

$$\begin{aligned} p_{x_t,\theta_t|Y_{t-1}}(x,\theta) &= \int_{\mathbb{R}^n} p_{x_t,\theta_t,x_{t-1}|Y_{t-1}}(x,\theta,x')dx' = \\ &= \int_{\mathbb{R}^n} p_{x_t|x_{t-1},\theta_t,Y_{t-1}}(x\,|\,x',\theta)\,p_{x_{t-1},\theta_t|Y_{t-1}}(x',\theta)dx' \end{aligned} \tag{9}$$

Because w_t is independent of Y_{t-1}, from (4) follows that x_t is conditionally independent of Y_{t-1} given (x_{t-1},θ_{t-1}). Hence (9) yields

$$p_{x_t,\theta_t|Y_{t-1}}(x,\theta) = \int_{\mathbb{R}^n} p_{x_t|x_{t-1},\theta_t}(x\,|\,x',\theta)\,p_{x_{t-1},\theta_t|Y_{t-1}}(x',\theta)dx' \tag{10}$$

Substitution of (8) into (10) yields

$$p_{x_t,\theta_t|Y_{t-1}}(x,\theta) = \int_{\mathbb{R}^n} p_{x_t|\theta_t,x_{t-1}}(x\,|\,\theta,x').\left[\sum_{\eta\in\mathbb{M}} \Pi_{\eta\theta}(x')\,p_{x_{t-1},\theta_{t-1}|,Y_{t-1}}(x',\eta)\right]dx' \tag{11}$$

Together with Bayes theorem for the measurement update step, (11) yields (12) in the Theorem below.

Theorem 1
Let $h(\theta,x,w)$ be such that the density $p_{y_t|x_t,\theta_t}(y\,|\,x,\theta)$ is measurable for all $(\theta,x)\in\mathbb{M}\times\mathbb{R}^n$, let (θ_0,x_0) admit a measurable density $p_{\theta_0,x_0}(\theta,x)$, and let the transition kernel of the Markov process $\{\theta_t,x_t\}$ admit a measurable transition density $p_{x_t,\theta_t|x_{t-1},\theta_{t-1}}(x,\theta\,|\,x',\eta)$, for all (η,x'), $(\theta,x)\in\mathbb{M}\times\mathbb{R}^n$, then the conditional density satisfies

$$\begin{aligned} p_{x_t,\theta_t|Y_t}(x,\theta) = p_{y_t|x_t,\theta_t}(y_t\,|\,x,\theta)\int_{\mathbb{R}^n} p_{x_t|\theta_t,x_{t-1}}(x\,|\,\theta,x') \\ .\sum_{\eta\in\mathbb{M}}\left[\Pi_{\eta\theta}(x')\,p_{x_{t-1},\theta_{t-1}|,Y_{t-1}}(x',\eta)\right]dx'/c_t \end{aligned} \tag{12}$$

with c_t a normalization constant.

Proof: (using Chapman-Kolmogorov)

Because $p_{y_t|x_t,\theta_t}(y\,|\,x,\theta)\in(0,\infty)$ for all $(y,\theta,x)\in\mathbb{R}^n\times\mathbb{M}\times\mathbb{R}^n$, application of (generalised) Bayes theorem (e.g. [43], p.129) yields:

$$p_{x_t,\theta_t|Y_t}(x,\theta) = p_{y_t|x_t,\theta_t}(y_t\,|\,x,\theta)\,p_{x_t,\theta_t|Y_{t-1}}(x,\theta)/c_t \tag{13}$$

Because $\{x_t,\theta_t\}$ is a Markov process, the evolution from $t-1$ to t satisfies the Chapman-Kolmogorov equation

$$p_{x_t,\theta_t|Y_{t-1}}(x,\theta) = \int_{\mathbb{R}^n}\sum_{\eta\in\mathbb{M}} p_{x_t,\theta_t|x_{t-1},\theta_{t-1}}(x,\theta\,|\,x',\eta)\,p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x',\eta)dx' \tag{14}$$

Factoring the Markov transition density in (14) yields

$$\begin{aligned} p_{x_t,\theta_t|x_{t-1},\theta_{t-1}}(x,\theta\,|\,x',\eta) &= p_{x_t|\theta_t,x_{t-1},\theta_{t-1}}(x\,|\,\theta,x',\eta)\,p_{\theta_t|x_{t-1},\theta_{t-1}}(\theta\,|\,x',\eta) \\ &= p_{x_t|\theta_t,x_{t-1},\theta_{t-1}}(x\,|\,\theta,x',\eta)\,\Pi_{\eta\theta}(x') \end{aligned} \tag{15}$$

Because w_t is independent of θ_{t-1}, from (1) follows that x_t is conditionally independent of θ_{t-1} given x_{t-1} and θ_t. Using this in (15) yields

$$p_{x_t,\theta_t|x_{t-1},\theta_{t-1}}(x,\theta\,|\,x',\eta) = p_{x_t|\theta_t,x_{t-1}}(x\,|\,\theta,x')\Pi_{\eta\theta}(x') \tag{16}$$

Substituting (16) into (14) and shifting summation yields

$$p_{x_t,\theta_t|Y_{t-1}}(x,\theta) = \int_{\mathbb{R}^n} p_{x_t|\theta_t,x_{t-1}}(x\,|\,\theta,x'). \cdot\left[\sum_{\eta\in\mathbb{M}} \Pi_{\eta\theta}(x')\,p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x',\eta)\right]dx' \tag{17}$$

Substituting this into (13) yields (12). Q.E.D.

Eq. (12) shows that state dependent mode switching probabilities have a multiplicative effect on the mode-switching-conditional evolution of the Euclidean valued state x_t. For hybrid stochastic processes this means there are two kinds of multiplication of conditional densities with other densities. One is due to Bayesian updating of measurements and the other is due to state dependent switching between modes

If $c(\theta,x,u)$ is x-invariant then $\{\theta_t\}$ is a Markov process , and the transition probability $\Pi_{\eta\theta}(x)$ is x-invariant. Then the term $\Pi_{\eta\theta}(x)$ in (12) can be shifted out of the integration over dx', and we get the Corollary below.

Corollary 1 (for Markov jump systems)
Let the assumptions of Theorem 1 hold true, and let $c(\theta,x,u)$ be x-invariant for $(\theta,x)\in \mathbb{M}\times\mathbb{R}^n$, then

$$p_{x_t,\theta_t|Y_t}(x,\theta) = p_{y_t|x_t,\theta_t}(y_t\,|\,x,\theta) \cdot\sum_{\eta\in\mathbb{M}}\left[\Pi_{\eta\theta}\int_{\mathbb{R}^n} p_{x_t|\theta_t,x_{t-1}}(x\,|\,\theta,x')p_{x_{t-1},\theta_{t-1}|,Y_{t-1}}(x',\eta)dx'\right]/c_t \tag{18}$$

with c_t a normalization constant

Comparison of (18), for Markov jump systems, with (12), for non-Markov jump systems, shows a relative small difference: $\Pi_{\eta\theta}$ is outside and inside the integration over $x'\in\mathbb{R}^n$ respectively. Because of this difference, under the non-Markov jump situation it is no longer possible to make explicit use of IMM's mixing probabilities $\mu_{t,\eta|\theta}$ in the derivation of the exact Bayesian filter recursion (12). This derivation path has been followed in [1] to get recursion (18).

Remark 1: If *c* is *x*-invariant, *a* and *h* are linear in (*x*,*w*) and (*x*,*v*) respectively, $p_{x_0|\theta_0,Y_0}(.\,|\,\theta)$ is Gaussian for all θ, then $p_{x_t|\theta_t,Y_t}(.\,|\,\theta)$ is a mixture of N^{t+1} Gaussian densities [44] – [45]. If we stick to all these conditions with the exception of the *x*-invariant *c*, then the appearance of the $\Pi_{\eta\theta}(x)$ in (12) destroys this exact Gaussian mixture solution. The best hope for some novel Gaussian mixture solution then is that $\Pi_{\eta\theta}(x)$ has a Gaussian shape (or the shape of a finite Gaussian mixture) for every η,θ.

For the IMM particle filter development it is relevant to decompose the exact recursive filter equation in Theorem 1 into a sequence of basic transitions. The following sequence of transitions defines such a decomposition:

$$p_{x_{t-1},\theta_{t-1}|Y_{t-1}}\begin{cases}\xrightarrow{\text{Mode Switching}} p_{\theta_t|Y_{t-1}}\\ \xrightarrow{\text{State Interaction}} p_{x_{t-1}|\theta_t,Y_{t-1}}\end{cases}$$

$$p_{x_{t-1},\theta_t|Y_{t-1}} \xrightarrow{\text{State Prediction}} p_{x_t,\theta_t|Y_{t-1}}$$

$$p_{x_t,\theta_t|Y_{t-1}} \xrightarrow{\text{Correction}} p_{x_t,\theta_t|Y_t}$$

The output of the first two transitions is integrated through the following equation:

$$p_{x_{t-1},\theta_t|Y_{t-1}}(x,\theta) = p_{x_{t-1}|\theta_t,Y_{t-1}}(x\,|\,\theta)\,p_{\theta_t|Y_{t-1}}(\theta) \tag{19}$$

We follow the above four transitions in developing the IMM particle filter, and do this in two steps. As first step we characterize the four transitions using results from section III.

Mode switching: The mode switching transition characterizes how the conditional mode probabilities evolve from $t-1$ to t. By law of total probability

$$p_{\theta_t|Y_{t-1}}(\theta) = \int_{\mathbb{R}^n} p_{x_{t-1},\theta_t|Y_{t-1}}(x,\theta)dx$$

Substitution of (8) and subsequent evaluation yields

$$p_{\theta_t|Y_{t-1}}(\theta) = \int_{\mathbb{R}^n}\sum_{\eta\in\mathbb{M}}\Pi_{\eta\theta}(x)p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\eta)dx = \\ = \sum_{\eta\in\mathbb{M}}\left[\int_{\mathbb{R}^n} p_{x_{t-1}|\theta_{t-1},Y_{t-1}}(x\,|\,\eta)\Pi_{\eta\theta}(x)dx\cdot p_{\theta_{t-1}|Y_{t-1}}(\eta)\right] \tag{20}$$

State interaction: The state interaction transition characterizes how the conditional state densities are being mixed under influence of mode switching. By dividing both left and right hand terms in (8) by $p_{\theta_t|Y_{t-1}}(\theta)$ we get for $p_{\theta_t|Y_{t-1}}(\theta) > 0$:

$$p_{x_{t-1}|\theta_t,Y_{t-1}}(x|\theta)=\sum_{\eta\in\mathbb{M}}\left[\Pi_{\eta\theta}(x)p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x|\eta)\right]/\,p_{\theta_t|Y_{t-1}}(\theta) \quad (21)$$

Eq. (21) represents the IMM interaction step at the level of the mode-conditional density. For Markov jump linear systems this form has been derived for the first time in [42] eq. (20). It is also derived in [46], section II, following another path of derivation. In later expositions on IMM, e.g. [47], p.461, [48], p.453, eq. (21) is derived in an approximate way only.

Remark 2: If the condition $p_{\theta_t|Y_{t-1}}(\theta)>0$ is not satisfied for one or more θ-values then a practical way of handling this is to skip evaluation of eq. (21) and assume an arbitrarily bounded density for the corresponding $p_{x_{t-1}|\theta_t,Y_{t-1}}(x|\theta)$'s.

<u>State prediction</u>: The state prediction transition characterizes the evolution of mode conditional state densities from $t-1$ to t. From (1) we get

$$p_{x_t|x_{t-1},\theta_t}(x|x',\theta)=\int_{[0,1]}\delta\left(x-a(\theta,x',w)\right)p_{w_t}(w)dw \quad (22)$$

Substituting this in (10) yields

$$p_{x_t,\theta_t|Y_{t-1}}(x,\theta)=\int_{\mathbb{R}^n}\int_{[0,1]}\delta\left(x-a(\theta,x',w)\right)p_{w_t}(w)dw\cdot$$
$$\cdot p_{x_{t-1},\theta_t|Y_{t-1}}\left(x',\theta\right)dx' \quad (23)$$

<u>Correction:</u> The correction transition is characterized by (13).

IV. IMM PARTICLE FILTER

Now we use the characterizations of the four transitions for the development of the IMM particle filter (IMMPF, see Table 3).

At moment $t-1$ the IMM particle filter starts with the set of weighted particles $\left\{x_{t-1}^{\theta,j},\mu_{t-1}^{\theta,j};\theta\in\mathbb{M},j\in\{1,..,S\}\right\}$, thus with a total number of $N_P=MS$ particles. This set of weighted particles spans the empirical density

$$\tilde{p}_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\theta)=\sum_{j=1}^{S}\mu_{t-1}^{\theta,j}\delta\left(x-x_{t-1}^{\theta,j}\right) \quad (24)$$

as an approximation of the exact density $p_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\theta)$.

<u>Mode Switching step of IMMPF</u>:
Substituting approximation (24) into (20) and subsequent evaluation yields:

$$\tilde{p}_{\theta_t|Y_{t-1}}(\theta)=\int_{\mathbb{R}^n}\sum_{\eta\in\mathbb{M}}\left[\Pi_{\eta\theta}(x)\sum_{j=1}^{S}\mu_{t-1}^{\eta,j}\delta\left(x-x_{t-1}^{\eta,j}\right)\right]dx=$$
$$=\sum_{\eta\in\mathbb{M}}\sum_{j=1}^{S}\Pi_{\eta\theta}\left(x_{t-1}^{\eta,j}\right)\mu_{t-1}^{\eta,j} \quad (25)$$

This forms the mode switching step of the IMM particle filter.

<u>Interaction resampling step of IMMPF</u>:
Next, substituting approximation (24) into (21) and subsequent evaluation yields:

$$\tilde{p}_{x_{t-1}|\theta_t,Y_{t-1}}(x|\theta)=\sum_{\eta\in\mathbb{M}}\left[\Pi_{\eta\theta}(x)\sum_{j=1}^{S}\mu_{t-1}^{\eta,j}\delta\left(x-x_{t-1}^{\eta,j}\right)\right]/\,\tilde{p}_{\theta_t|Y_{t-1}}(\theta)=$$
$$=\sum_{\eta\in\mathbb{M}}\sum_{j=1}^{S}\Pi_{\eta\theta}\left(x_{t-1}^{\eta,j}\right)\mu_{t-1}^{\eta,j}\delta\left(x-x_{t-1}^{\eta,j}\right)/\,\tilde{p}_{\theta_t|Y_{t-1}}(\theta) \quad (26)$$

Eq. (26) makes clear that $\tilde{p}_{x_{t-1}|\theta_t,Y_{t-1}}(x|\theta)$ is an empirical density spanned by $N_P=MS$ particles, whereas (24) shows that $\tilde{p}_{x_{t-1}|\theta_{t-1},Y_{t-1}}(x|\theta)$ was an empirical density spanned by S particles only. A logical way to reduce the number of particles to span $\tilde{p}_{x_{t-1}|\theta_t,Y_{t-1}}(x|\theta)$ is to perform a resampling of eq. (26). This forms the interaction resampling step of the IMM particle filter.

Remark 3: Choosing the moment of interaction as the right moment of resampling is similar to in IMM doing the hypothesis merging step in combination with the interaction step rather than after measurement update [46].

<u>Prediction step of IMMPF</u>
Substituting the resampled version of $\tilde{p}_{x_{t-1}|\theta_t,Y_{t-1}}(x|\theta)$ and $\tilde{p}_{w_t}(w)=\delta\left(w-w_t^{\theta,j}\right)$ into (22) and subsequent evaluation yields:

$$\tilde{p}_{x_t,\theta_t|Y_{t-1}}(x,\theta)=\int_{\mathbb{R}^n}\int_{\mathbb{R}^{n'}}\delta\left(x-a(\theta,x',w)\right)\delta\left(w-w_t^{\theta,j}\right)dw\cdot$$
$$\cdot\sum_{j=1}^{S}\bar{\mu}_t^{\theta,j}\delta\left(x'-\bar{x}_{t-1}^{\theta,j}\right)=$$
$$=\sum_{j=1}^{S}\bar{\mu}_t^{\theta,j}\delta\left(x-a\left(\theta,\bar{x}_{t-1}^{\theta,j},w_{t-1}^{\theta,j}\right)\right)=$$
$$=\sum_{j=1}^{S}\bar{\mu}_t^{\theta,j}\delta\left(x-x_t^{\theta,j}\right) \quad (27.a)$$

with

$$x_t^{\theta,j}=a\left(\theta,\bar{x}_{t-1}^{\theta,j},w_t^{\theta,j}\right) \quad (27.b)$$

This forms the prediction step of the IMM particle filter.

Correction step of IMMPF

Substituting eq. (27) into (13) and subsequent evaluation yields

$$\tilde{p}_{x_t,\theta_t|Y_t}(x,\theta) = p_{y_t|x_t\theta_t}(y\,|\,x,\theta)\sum_{j=1}^{S}\bar{\mu}_t^{\theta,j}\delta\left(x - x_t^{\theta,j}\right)/c_t =$$

$$= \sum_{j=1}^{S}\mu_t^{\theta,j}\delta\left(x - x_t^{\theta,j}\right)/c_t \qquad (28.a)$$

with $\mu_t^{\theta,j} = \bar{\mu}_t^{\theta,j}\, p_{y_t|x_t,\theta_t}(y\,|\,x_t^{\theta,j},\theta)/c_t$ (28.b)

This forms the correction step of the IMM particle filter.

Output step of IMMPF

Because we have particle values for x_t only, we next determine the output equations of the IMM particle filter. From the law of total probability we have:

$$\tilde{p}_{\theta_t|Y_t}(\theta) = \int_{\mathbb{R}^n}\tilde{p}_{x_t,\theta_t|Y_t}(x,\theta)dx \qquad (29)$$

where $\tilde{p}_{\theta_t|Y_t}(\theta)$ is an approximation of $p_{\theta_t|Y_t}(\theta)$.

Substitution of (28.a) into (29) and subsequent evaluation yields

$$\tilde{p}_{\theta_t|Y_t}(\theta) = \int_{\mathbb{R}^n}\sum_{j=1}^{S}\mu_t^{\theta,j}\delta\left(x - x_t^{\theta,j}\right)dx =$$

$$= \sum_{j=1}^{S}\mu_t^{\theta,j} \qquad (30)$$

Through dividing the left and right hand terms of (28.a) by $\tilde{p}_{\theta_t|Y_t}(\theta)$ and subsequent evaluation we get for $\tilde{p}_{\theta_t|Y_t}(\theta) > 0$:

$$\tilde{p}_{x_t|\theta_t,Y_t}(x\,|\,\theta) = \sum_{j=1}^{S}\mu_t^{\theta,j}\delta\left(x - x_t^{\theta,j}\right)/\,\tilde{p}_{\theta_t|Y_t}(\theta) \qquad (31)$$

Substitution of (30) into (31) yields:

$$\tilde{p}_{x_t|\theta_t,Y_{t-1}}(x\,|\,\theta) = \sum_{j=1}^{S}\mu_t^{\theta,j}\delta\left(x - x_t^{\theta,j}\right)/\sum_{j=1}^{S}\mu_t^{\theta,j} \qquad (32)$$

Equations (30) and (32) characterize the output step of the IMM particle filter.

Remark 4: If $y = h(\theta,x,v)$ has for each θ, x an inverse $v = g(\theta,x,y)$ which is differentiable in y, then (e.g. Kitagawa, 1996):

$$p_{y_t|x_t,\theta_t}(y\,|\,x,\theta) = p_{v_t}(g(\theta,x,y))\left[\frac{\partial g(\theta,x,y)}{\partial y}\right]$$

The main changes of the IMMPF over the SIR PF are:

- Resample fixed number of particles per mode;
- Probabilities for $\{\theta_t\}$ instead of particles for $\{\theta_t\}$;
- Resampling after interaction/mixing rather than after measurement update.

Table 1. SIR Particle Filter (SIR PF) cycle

SIR	$\tilde{p}_{x_{t-1},\theta_{t-1}\|Y_{t-1}} \to \tilde{p}_{x_t,\theta_t\|Y_t}$
Particles $\left\{\mu_{t-1}^j \in [0,1], \theta_{t-1}^j \in \mathbb{M}, x_{t-1}^j \in \mathbb{R}^n;\ j = 1,...,N_p\right\}$ $\tilde{p}_{x_{t-1},\theta_{t-1}\|Y_{t-1}}(x,\theta) = \sum_{j=1}^{N_p}\mu_{t-1}^j\chi(\theta,\theta_{t-1}^j)\delta(x - x_{t-1}^j)$	
For j=1,..., N_p: • Generate w_t^j and u_t^j i.i.d. from $p_{w_t}(w)$ and $p_{u_t}(u)$ • Evolution: $\bar{\theta}_t^j = c\left(\theta_{t-1}^j, x_{t-1}^j, u_t^j\right)$ $\bar{x}_t^j = a\left(\theta_t^j, x_{t-1}^j, w_t^j\right)$ • Correction: $\bar{\mu}_t^j = \mu_{t-1}^j \cdot N_m\left\{y_t; h(\bar{\theta}_t^j,\bar{x}_t^j), g(\bar{\theta}_t^j,\bar{x}_t^j)g(\bar{\theta}_t^j,\bar{x}_t^j)^T\right\}/c_t$ with c_t such that $\sum_{j=1}^{N_p}\bar{\mu}_t^j = 1$ • Resampling: $\mu_t^j = 1/N_p$ $\left(x_t^j, \theta_t^j\right) \sim \sum_{j=1}^{N_p}\bar{\mu}_t^j\chi\left(\theta,\bar{\theta}_t^j\right)\delta(x - \bar{x}_t^j)$	

Table 2. IMM Particle Filter (IMMPF) cycle

IMM Particle Filter (IMMPF) cycle

Particles

$$\left\{\mu_{t-1}^{\theta,j}\in[0,1],\, x_{t-1}^{\theta,j}\in\mathbb{R}^n,\theta\in\mathbb{M};\, j=1,\dots,N_p/M\right\}$$

$$\tilde{p}_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\theta)=\sum_{j}^{N_p/M}\mu_{t-1}^{\theta,j}\delta\left(x-x_{t-1}^{\theta,j}\right)$$

- Mode switching:

$$\bar{\gamma}_t(\theta)\triangleq\tilde{p}_{\theta_t|Y_{t-1}}(\theta)=\sum_{\eta\in\mathbb{M}}\sum_{j=1}^{N_p/M}\Pi_{\eta\theta}\left(x_{t-1}^{\eta,j}\right)\mu_{t-1}^{\eta,j}$$

- Interaction resampling:

$$\bar{\mu}_t^{\theta,j}=\bar{\gamma}_t(\theta)M/N_p,\quad j\in\left\{1,\dots,N_p/M\right\},\ \theta\in\mathbb{M}$$

If $\bar{\gamma}_t(\theta)=0$ then $\bar{x}_{t-1}^{\theta,j}=x_{t-1}^{\theta,j}$, otherwise

$$\bar{x}_{t-1}^{\theta,j}\sim\sum_{\eta\in\mathbb{M}}\sum_{j=1}^{N_p/M}\Pi_{\eta\theta}\left(x_{t-1}^{\eta,j}\right)\mu_{t-1}^{\eta,j}\delta\left(x-x_{t-1}^{\eta,j}\right)/\bar{\gamma}_t(\theta)$$

- Prediction:

$w_t^{\theta,j}\sim p_{w_t}(w)$ i.i.d. for $(\theta,j)\in\mathbb{M}\times\{1,\dots,N_p/M\}$

$$x_t^{\theta,j}=a\left(\theta,\bar{x}_{t-1}^{\theta,j},w_t^{\theta,j}\right)$$

- Correction:

$$\mu_t^{\theta,j}=\bar{\mu}_t^{\theta,j}\cdot p_{y_t|x_t,\theta_t}(y_t\,|\,x_t^{\theta,j},\theta)/c_t$$

with c_t such that $\sum_{j=1}^{N_p/M}\sum_{\theta\in\mathbb{M}}\mu_t^{\theta,j}=1$

- Output:

$$\gamma_t(\theta)\triangleq\tilde{p}_{\theta_t|Y_t}(\theta)=\sum_{j=1}^{N_p/M}\mu_t^{\theta,j}$$

$$\tilde{p}_{x_t|\theta_t,Y_t}(x\,|\,\theta)=\sum_{j=1}^{N_p/M}\mu_t^{\theta,j}\delta\left(x-x_t^{\theta,j}\right)/\gamma_t(\theta)\ \text{ if }\ \gamma_t(\theta)>0$$

Table 3. Comparison of particle filter characteristics

	SIR PF (table 1)	HPF (table 4)	IMMPF (table 2)
Memorize θ_t-values	Yes	No	No
Mode-switching	Simulation	Simulation	Analytical
x_t-Prediction	Simulation	Simulation	Simulation
Correction	Standard	Standard	Standard
Resampling timing	After correction	After correction	Combined with Interaction
Resampling type	Equal weights	Fixed number of particles per mode	Fixed number of particles per mode

Table 4. Hybrid Particle Filter (HPF) cycle

Hybrid Particle Filter (HPF) cycle

Particles

$$\left\{\mu_{t-1}^{\theta,j}\in[0,1],\, x_{t-1}^{\theta,j}\in\mathbb{R}^n,\theta\in\mathbb{M};\, j=1,\dots,N_p/M\right\}$$

$$\tilde{p}_{x_{t-1},\theta_{t-1}|Y_{t-1}}(x,\theta)=\sum_{j=1}^{N_p/M}\mu_{t-1}^{\theta,j}\delta\left(x-x_{t-1}^{\theta,j}\right)$$

- Mode switching:

$$u_{t-1}^{\theta,j}\sim p_{u_t}(u)$$

$$\bar{\theta}_t^{\theta,j}=c\left(\theta,x_{t-1}^{\theta,j},u_t^{\theta,j}\right)$$

- Prediction:

$w_t^{\theta,j}\sim p_{w_t}(w)$ i.i.d., $\theta\in\mathbb{M}, j\in\{1,.,N_p/M\}$

$$\bar{x}_t^{\theta,j}=a\left(\bar{\theta}_t^{\theta,j},x_{t-1}^{\theta,j},w_t^{\theta,j}\right)$$

- Correction:

$$\mu_t^{\theta,j}=\mu_{t-1}^{\theta,j}\cdot p_{y_t|x_t,\theta_t}(y_t\,|\,\bar{x}_t^{\theta,j},\bar{\theta}_t^{\theta,j})/c_t$$

with c_t such that $\sum_{j=1}^{N_p/M}\sum_{\theta\in\mathbb{M}}\mu_t^{\theta,j}=1$

- Resampling:

$$\gamma_t(\theta)=\sum_{j=1}^{N_p/M}\sum_{\eta\in\mathbb{M}}\mu_t^{\theta,j}\chi\left(\bar{\theta}_t^{\eta,j},\theta\right)$$

$$\mu_t^{\theta,j}=\gamma_t(\theta)M/N_p$$

$$x_t^{\theta,j}\sim\sum_{j=1}^{N_p/M}\sum_{\eta\in\mathbb{M}}\mu_t^{\theta,j}\delta\left(\bar{\theta}_t^{\theta,j},\theta\right)\delta\left(x-\bar{x}_t^{\theta,j}\right)/\gamma_t(\theta)$$

i.i.d. for $(\theta,j)\in\mathbb{M}\times\left\{1,\dots,N_p/M\right\}$ if $\gamma_t(\theta)>0$

Table 3 provides an overview of similarities and differences between SIR PF applied to stochastic hybrid systems and IMM PF. In addition, table 3 extends this comparison to a version of the SIR particle filter that has shown to work well for multi-target tracking [14]. We refer to this scheme (see Table 4) as Hybrid Particle Filter (HPF).

V. MONTE CARLO SIMULATIONS

In this section some Monte Carlo simulation results are given for the IMM Particle Filter (IMMPF), the standard Particle Filter (PF) and the IMM algorithm. In addition we also give simulation results for a Hybrid Particle Filter (HPF) which differs from the standard PF by resampling a fixed number of particles per mode. For each of the particle filters we used a total of N_p=10000 and N_p=1000 particles respectively. The simulations primarily aim at gaining insight in the behavior and performance of the filters in case of less frequent switching. In the example scenarios there is an object moving with two possible modes, i.e. $M = 2$. One mode is constant velocity and the other mode is constant acceleration. The object starts with zero velocity and continues this for 40 scans. After scan 40 the object starts to accelerate with at a value equal to the standard deviation σ_a of acceleration values. In scenarios 1 and 2 the object continues with constant velocity after scan 60, while in scenarios 3 and 4 the object continues accelerating. In each simulation, the filters start with perfect estimates and run for 100 scans.

The model considered is a Markov jump linear system:

$$x_t = A(\theta_t)x_{t-1} + B(\theta_t)w_t$$
$$y_t = [1\ 0\ 0]x_t + \sigma_m v_t$$

with $\mathbb{M} = \{1, 2\}$ and

$$A(1) = \begin{bmatrix} 1 & T_s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad A(2) = \begin{bmatrix} 1 & T_s & \frac{1}{2}T_s^2 \\ 0 & 1 & T_s \\ 0 & 0 & \alpha \end{bmatrix}$$

$$B(1) = \sigma_a \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad B(2) = \sigma_a \begin{bmatrix} 0 \\ 0 \\ \sqrt{1-\alpha^2} \end{bmatrix}$$

$$\Pi = \begin{bmatrix} 1 - \frac{T_s}{\tau_1} & \frac{T_s}{\tau_1} \\ \frac{T_s}{\tau_2} & 1 - \frac{T_s}{\tau_2} \end{bmatrix}$$

where σ_a represents the standard deviation of acceleration noise, σ_m represents the standard deviation of the measurement error, T_s is the time duration between two successive observation moments $t-1$ and t, τ_1 and τ_2 are the mean durations of modes 1 and 2 respectively, the parameter $\alpha \in (0,1]$ allows the acceleration in mode 2 to vary randomly in time. Table I gives the scenario parameter values that are being used for the Monte Carlo simulations.

For each of the scenarios Monte Carlo simulations containing 100 runs have been performed for each of the filters. To make the comparison more meaningful, for all filters the same random number streams were used. The results of the Monte Carlo simulations of the four scenarios are shown as follows:

– The position RMS errors in figures 1 through 4.
– The speed RMS errors in figures 5 through 8.
– The acceleration RMS errors in figures 9 through 12
– The computational load in Table II.

TABLE I
SCENARIO PARAMETER VALUES

Scenario	α	σ_a [m/s^2]	σ_m [m]	τ_1 [s]	τ_2 [s]	T_S [s]
1	0.9	50	30	50	5	1
2	0.9	50	30	5000	5	1
3	0.9	1	30	50	5	1
4	0.9	1	30	5000	500	1

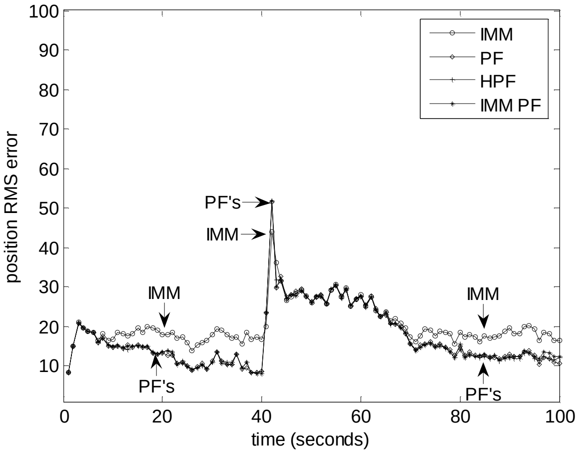
TABLE II
COMPUTATIONAL LOAD PER SCAN (10^{-3} s)

N_p	IMM	PF	HPF	IMMPF
10^4	4	138	115	96
10^3	4	19	13	11

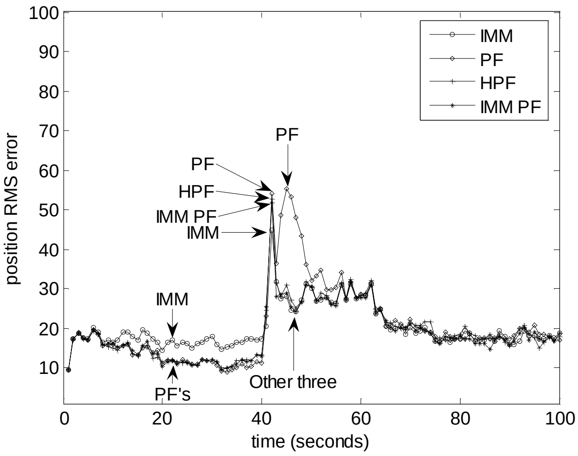
Scenario 1: In this scenario, the target accelerates with 5g between 40 s and 60 s. The model used by the particle filters expect accelerations to happen about once per minute. With N_p=10^4 particles, all three particle filters perform similarly well; they converge to a lower value during uniform motion than IMM does. As a side effect, the peak RMS error at the start of acceleration is for the particle filters slightly higher than it is for IMM. These results agree well with those in [6] – [7]. Reduction of the number of particles to N_p=10^3 affects PF dramatically, but has negligible impact on HPF and IMMPF.

Scenario 2:
In this scenario, the target accelerates with 5g between 40 s and 60 s. The model used by the particle filters expect accelerations to happen less than once per hour.
With N_p=10^4 particles, IMMPF performs marginally better than IMM does, while PF performs dramatically worse. HPF performs significantly worse during the initial acceleration period only. Reduction of the number of particles to N_p=10^3 has a negative effect on the convergence during constant velocity for all three particle filters. Moreover, during the period of acceleration, PF and HPF worsen dramatically in performance.

Scenario 3:
The target accelerates with 0.1g after 40 s. The model used by the particle filters expect accelerations to happen about once per minute. With N_p=10^4 particles, all three particle filters perform equally well, and significantly better than IMM does. Reduction of the number of particles to N_p=10^3 has a clear negative effect for the standard PF, but does not affect IMMPF and HPF.

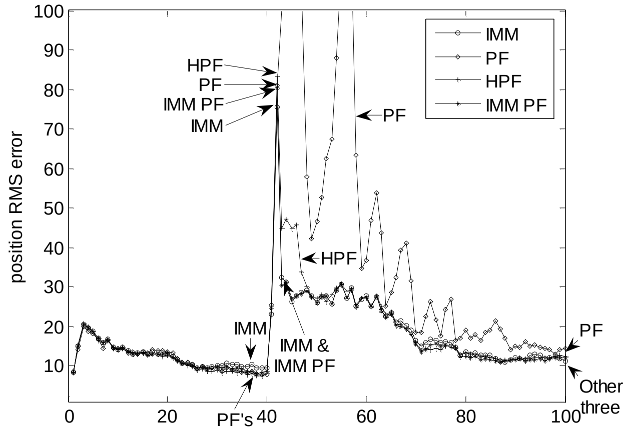


a. N_p=10^4 particles

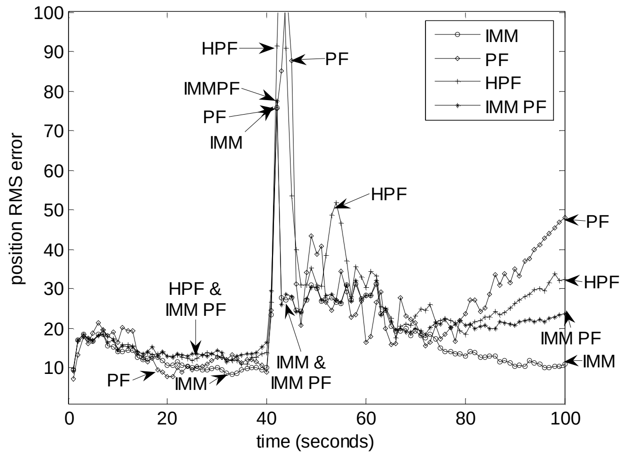


b. N_p=10^3 particles

Figure 1. Scenario 1. The target accelerates with 5g between 40 s and 60 s. The particle filter parameters are $\sigma_a = 5\text{g}$, $\tau_1 = 50\text{s}$ and $\tau_2 = 5\text{s}$.



a. N_p=10^4 particles



b. N_p=10^3 particles

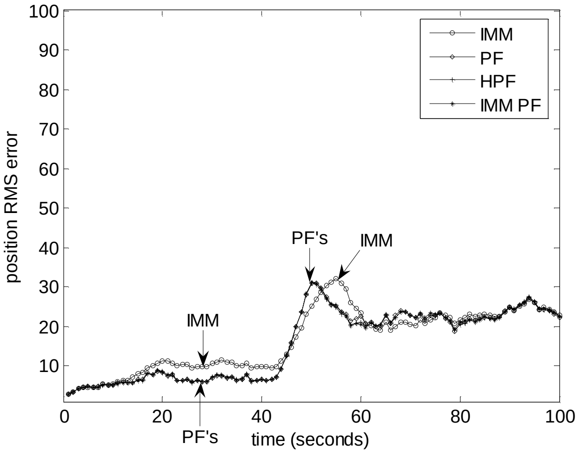
Figure 2. Scenario 2. The target accelerates with 5g between 40 s and 60 s. The particle filter parameters are $\sigma_a = 5\text{g}$, $\tau_1 = 5000\text{s}$ and $\tau_2 = 5\text{s}$.

Scenario 4:

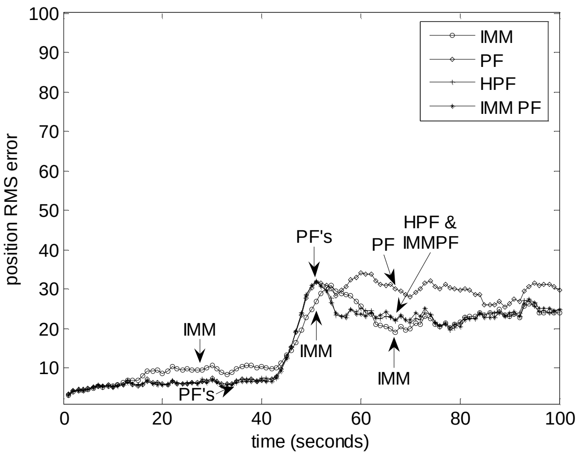
The target accelerates with 0.1 g after 40 s. The model used by the particle filters expect accelerations to happen less than once per hour. With N_p=10^4 particles, all four filters, except the standard PF, perform similarly well. The standard PF performs dramatically worse during constant acceleration. Reduction of the number of particles to 10^3 has a clear negative effect for the standard PF and the HPF, but not for the IMMPF.

Summary of Monte Carlo simulation results:

With N_p=10^4 particles, all three particle filters perform better than IMM for scenarios 1 and 3. For scenarios 2 and 4 however, IMM and IMMPF perform similarly well, while the standard PF performs less good on sudden acceleration, and the HPF response is less good for acceleration in scenario 2 only. With N_p=10^3 particles the performance of PF degrades for all scenarios, of HPF for scenarios 2 and 4, and of IMMPF for scenario 2 only.

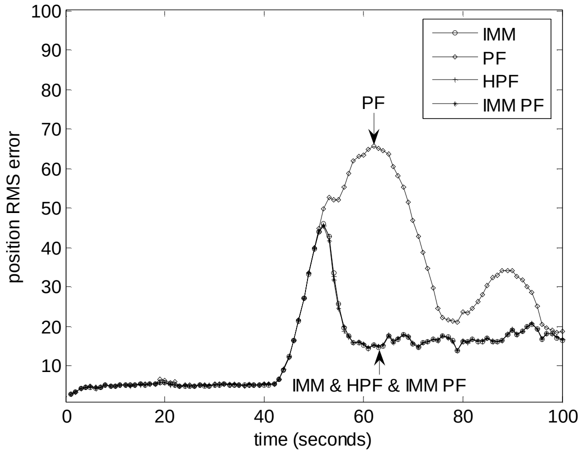


a. N_p=10^4 particles

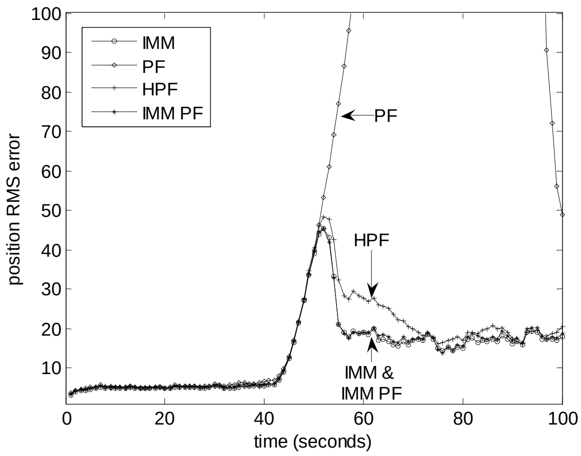


b. N_p=10^3 particles

Figure 3. Scenario 3. The target accelerates with 0.1g after 40 s. The particle filter parameters are $\sigma_a = 0.1\text{g}$, $\tau_1 = 50\text{s}$ and $\tau_2 = 5\text{s}$.



a. N_p=10^4 particles



b. N_p=10^3 particles

Figure 4. Scenario 4. The target accelerates with 0.1 g after 40 s. The particle filter parameters are $\sigma_a = 0.1\text{g}$, $\tau_1 = 5000\text{s}$ and $\tau_2 = 500\text{s}$.

VI. CONCLUDING REMARKS

In this paper we considered filtering of stochastic hybrid systems that go beyond the well known Markov jump system. We derived the exact Bayesian filter and used this for the development of a novel particle filter for discrete time stochastic hybrid systems. Because of its similarity with the interaction step of IMM, this novel particle filter is referred to as IMM particle filter (IMMPF). Through MC simulations for four scenarios, IMMPF has been tested and compared with standard PF and IMM. With 10^4 particles, the IMMPF performs well for all four scenarios. The computational load is 25 times the load of IMM. The computational load of the standard PF is even higher. As expected, the IMMPF works well for all four scenarios including ones where standard PF or IMM has problems. Hence IMMPF is the preferred particle filter for stochastic hybrid systems.

Next we compare IMMPF with IMM. For the regular switching scenarios 1 and 3, the IMMPF has some performance advantage over IMM, also when the number of particles is down to 10^3. The computational load of IMMPF is then three times higher than the load of IMM. However, for the scenarios with infrequent mode switching rate (scenarios 2 and 4), the IMM performs similarly well as the IMMPF. In this case the main advantage over IMM is that IMMPF incorporates various kinds of deviations from the Markov jump linear IMM (i.e. non-linear and non-Markov mode switching).

Because IMMPF uses particles for the conditional densities given the mode, it is likely that several of the SIR improvement methods referred to in the introduction of this paper, combine quite well with IMMPF, and are expected to allow a significant reduction of the number of particles. This is topic of research in a follow-up study.

ACKNOWLEDGEMENT

The authors would like to thank anonymous reviewers for their valuable suggestions in improving the paper.

REFERENCES

[1] Y. Bar-Shalom, S. Challa, H.A.P. Blom, IMM estimator versus optimal estimator for hybrid systems, IEEE Tr. Aerospace and Electronic Systems, Vol. 41 (2005), pp.986-991.

[2] V.J. Mathews, J.K. Tugnait, Detection and estimation with fixed lag for abruptly changing systems, IEEE Tr. Aerospace and Electronic Systems, Vol. 19 (1983), pp. 730-739.

[3] L. Campo, P. Mookerjee, Y. Bar-Shalom, State Estimation for Systems with a Sojourn-Time-Dependent Markov Switching Model, IEEE Tr. Automatic Control, Vol. 36 (1991), pp. 238-243.

[4] H.A.P. Blom, Hybrid state estimation for systems with semi-Markov switching coefficients, Proc. 1st European Control Conf., Grenoble, 1991, pp. 1132-1137.

[5] X. Koutsoukis, J. Kurien, F. Zhao, "Monitoring and Diagnosis of hybrid systems using particle filtering methods," Proc. Mathematical Theory of Networks and Systems (MTNS), 2002.

[6] S. McGinnity, G. W. Irwin, "Multiple Model Bootstrap Filter for Maneuvering Target Tracking," *IEEE Tr. on Aerospace and Electronic System*s, Vol. 36, 2000, pp. 1006-1012.

[7] S. McGinnity, G. W. Irwin, "Maneuvering Target Tracking using a Multiple-Model Bootstrap Filter," Eds. A. Doucet, N. de Freitas and N. Gordon, *Sequential Monte Carlo Methods in Practic*e, Springer 2001, pp. 479-497.

[8] D.S. Angelova, T.A. Semerdjiev, V.P. Jilkov, E.A. Semerdjiev, Application of Monte Carlo method for tracking maneuvering target in clutter, Mathematics and Computers in Simulation, Vol. 1851 (2000), pp. 1-9.

[9] R. Chen, J.S. Liu, Mixture Kalman filters, J. Royal Statist. Soc. B, Vol. 62, Part 3 (2000), pp. 493-508.

[10] A. Doucet, A. Logothetis, V. Krishnamurthy, Stochastic sampling algorithms for state estimation of jump Markov linear systems, IEEE Tr. Automatic Control, Vol. 45 (2000), pp. 188-202.

[11] A. Doucet, N. J. Gordon, V. Krishnamurthy, "Particle Filters for State Estimation of Jump Markov Linear Systems," *IEEE Tr. on Signal Processin*g, Vol. 49, 2001, pp. 613-624.

[12] C. Musso, N. Oudjane, F. Le Gland, "Improving Regularised Particle Filters," Eds. A. Doucet, N. de Freitas and N. Gordon, *Sequential Monte Carlo Methods in Practic*e, Springer 2001, pp. 247-271.

[13] C. Andrieu, M. Davy, A. Doucet, Efficient particle filtering for jump Markov systems. Application to time-varying autoregressions, IEEE Tr. Signal Processing, Vol. 51 (2003), pp. 1762-1770.

[14] H.A.P. Blom, E.A. Bloem, "Joint IMMPDA Particle filter," Proc. 6th Int. Conf. on Information Fusion, 2003, pp. 785-792.

[15] Y. Boers, H. Driessen, Hybrid state estimation: a target tracking application, Automatica, Vol. 38 (2002), pp. 2153-2158.

[16] Y. Boers, J.N. Driessen, Interacting multiple model particle filter, IEE Proc. Radar Sonar Navig., Vol. 150 (2003), pp. 344-349.

[17] H. Driessen, Y. Boers, An Efficient Particle Filter for Jump Markov Nonlinear Systems, Proc. IEE Colloquium on Target Tracking, March 2004.

[18] B. Ristic, S. Arulampalam, N. Gordon, Beyond the Kalman filter - Particle filters for tracking applications, Artech House, 2004.

[19] M. Morelande, S. Challa, Maneuvering target tracking in clutter using particle filters, IEEE Tr. AES, Vol. 41 (2005), pp. 252-270.

[20] D.B. Rubin, "Using the SIR algorithm to simulate posterior distributions," Eds. J.M. Bernardo et al., Bayesian Statistics 3, Oxford Univ. Press, 1988, pp. 395-402.

[21] N.J. Gordon, D. J. Salmond, A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings*-F, Vol. 140, pp. 107-113, 1993.

[22] P. Del Moral, Nonlinear filtering: interacting particle solution, Markov Processes and Related Fields, Vol. 2 (1996), pp. 555-580.

[23] G. Kitagawa, Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, J. Computational and Graphical Statistics, Vol. 5 (1996), pp. 1-25.

[24] A. Doucet, "On sequential simulation-based methods for Bayesian filtering," Technical report CUED / F-INFENG / TR-310, Univ. of Cambridge, UK, 1998.

[25] P. Del Moral, A. Guionnet, Large deviation for interacting particle systems - Application to nonlinear filtering problems, Stochastic Processes and their Applications, Vol. 78 (1998), pp. 69-95.

[26] D. Crisan, A. Doucet, A survey of convergence results on particle filtering methods for practitioners, IEEE Tr. Signal Processinbg, Vol. 50 (2002), pp. 736-746.

[27] F.E. Daum, J. Huang, The curse of dimensionality for particle filters, Proc. IEEE Conf. on Aerospace, Big Sky, MT, March 2003.

[28] M. Isard, A. Blake, Condensation – Conditional density propagation for visual tracking, J. of Computer Vision, Vol. 29 (1998), pp. 5-28.

[29] Y. Boers, J.N. Driessen, F. Verschure, W.P.M.H. Heemels, A. Juloski, A multi target track before detect application, Proc.IEEE CVPR Conf, Workshop on multi object tracking, Madison, WI, June 2003.

[30] J.S. Liu, R. Chen, Sequential Monte Carlo methods for dynamical systems, J. American Statistical Association, Vol. 93 (1998), pp. 1032-1044.

[31] W.R. Gilks, C. Berzuini, Following a moving target – Monte Carlo inference for dynamic Bayesian models, J. Royal Statistical Society B, Vol. 63 (2001), pp. 127-146.

[32] A. Doucet, S. Godsill, C. Andrieu, On sequential Monte Carlo sampling methods for Bayesian filtering, Statistics and Computing, Vol. 10 (2000), pp. 197-208.

[33] M.K. Pitt, N. Shephard, Filtering via simulation: auxiliary particle filters, J. of American Statistical Association, Vol. 94 (1999), pp. 590-599.

[34] M. Hürzeler, H.R. Künsch, Monte Carlo approximations for general state-space models, J. Computational and Graphical Statistics, Vol. 7 (1998), pp. 175-193.

[35] M.S. Arulampam, S. Maskell, N. Gordon, T. Clapp, A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, IEEE Tr. Sgnal Processing, Vol. 50 (2002), pp. 174-188.

[36] Y. Rui, Y. Chen, Better proposal distributions: object tracking using unscented particle filter, Proc. IEEE 2001, pp. II-786-793.

[37] G. Casella, C. Robert, Rao-Blackwellisation of sampling schemes, Biometrika, Vol. 83 (1996), pp. 81-94.

[38] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, P.-J. Nordlund, Particle filters for positioning, navigation and tracking, IEEE Tr. SP, Vol. 50 (2002), pp. 425-437.

[39] S. Maskell, M. Rollason, N. Gordon, D. Salmond, Efficient particle filtering for multiple target tracking with application to tracking in structured images, Proc. SPIE, Signal and Data Processing of Small Targets, Vol. 4728, April 2002.

[40] M.R. Morelande, S. Challa, An algorithm for tracking group targets, Proc. Workshop on Multiple Hypothesis Tracking: A tribute to S. Blackman, San Diego, May 2003.

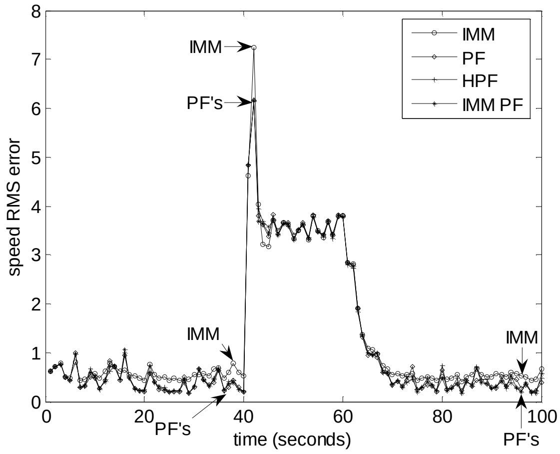
[41] H.A.P. Blom, E.A. Bloem, "Particle filtering for stochastic hybrid systems," Proc. IEEE CDC04, Bahamas, December 2004.

[42] H.A.P. Blom, "An efficient filter for abruptly changing systems," *Proc. of the 23rd IEEE CDC*, 1984, pp.656-658.

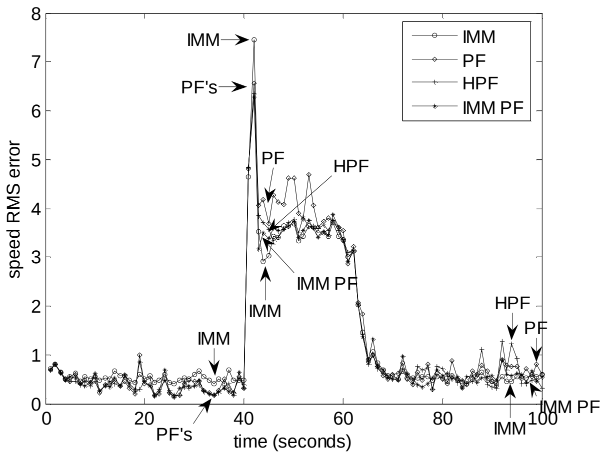
[43] J.M. Bernardo, A.F.M. Smith, Bayesian theory, Wiley, 1994.

[44] J.K. Tugnait, A.H. Haddad, "A detection-estimation scheme for state estimation in switching environments," Automatica, Vol. 15 (1979), pp. 477-481.

[45] R.J. Elliott, F. Dufour, D.D. Sworder, Exact hybrid filters in discrete time, IEEE Tr. AC, Vol. 41 (1996), pp. 1807-1810.
[46] H.A.P. Blom, Y. Bar-Shalom, "The Interacting Multiple Model algorithm for systems with Markovian switching coefficients," *IEEE Tr. on Automatic Control*, Vol. 33 (1988), pp. 780-783.
[47] Y. Bar-Shalom, X. R. Li, Estimation and tracking: principles, techniques and software, Artech House, 1993.
[48] Y. Bar-Shalom, X.R. Li, T. Kirubarajan, Estimation, tracking and navigation, Wiley-Interscience, New York, 2001.

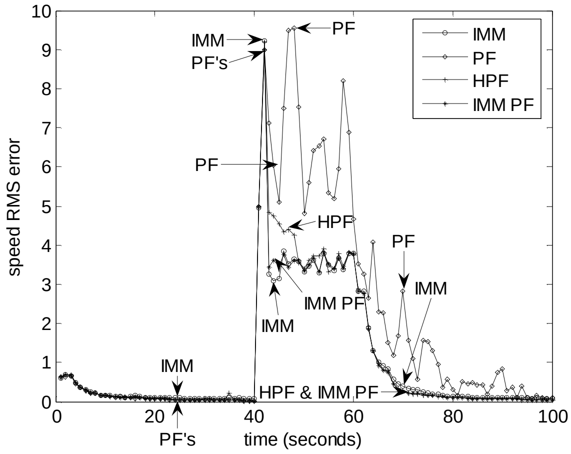


a. 10^4 particles

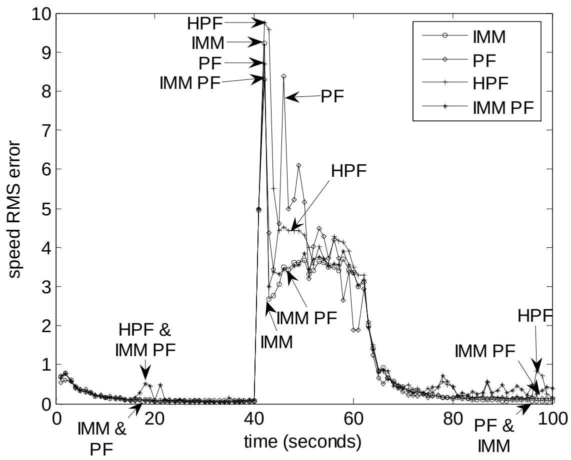


b. 10^3 particles

Figure 5. Scenario 1. The target accelerates with 5g between 40 s and 60 s. The particle filter parameters are $\sigma_a = 5\text{g}$, $\tau_1 = 50\text{s}$ and $\tau_2 = 5\text{s}$.

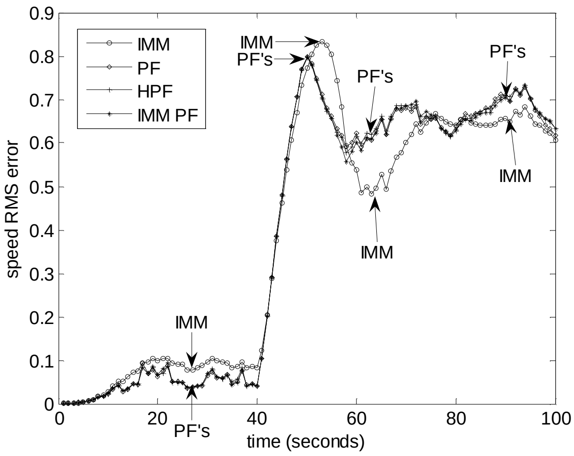


a. 10^4 particles

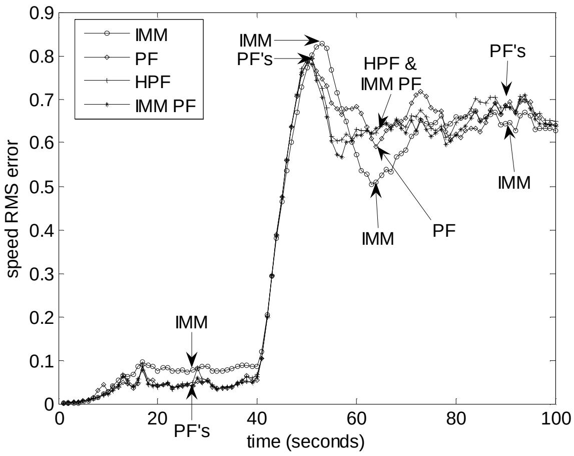


b. 10^3 particles

Figure 6. Scenario 2. The target accelerates with 5g between 40 s and 60 s. The particle filter parameters are $\sigma_a = 5\text{g}$, $\tau_1 = 5000\text{s}$ and $\tau_2 = 5\text{s}$.

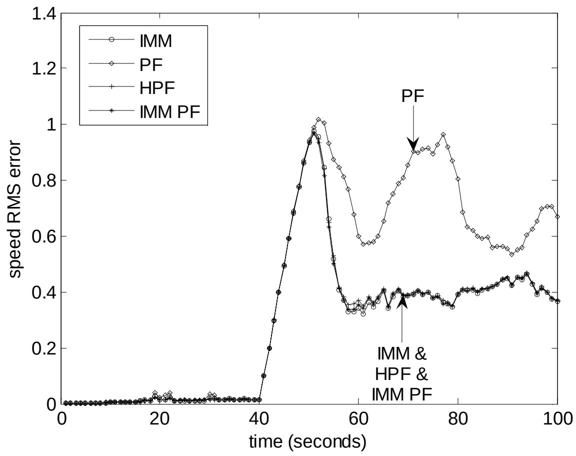


a. 10^4 particles

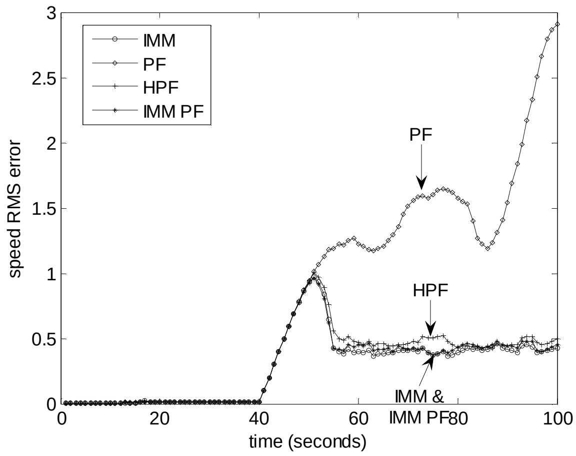


b. 10^3 particles

Figure7. Scenario 3. The target accelerates with 0.1 g after 40 s. The particle filter parameters are $\sigma_a = 0.1\text{g}$, $\tau_1 = 50\text{s}$ and $\tau_2 = 5\text{s}$.

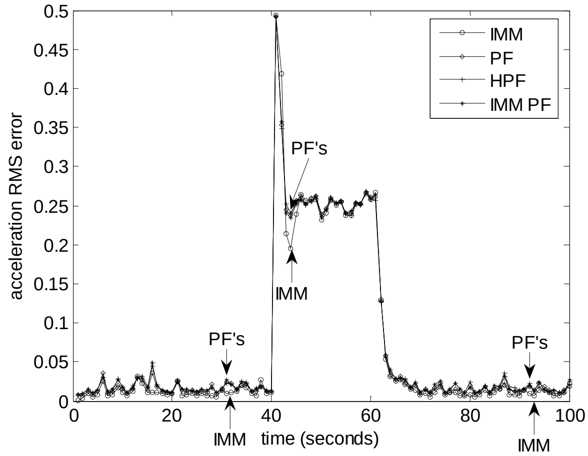


a. 10^4 particles

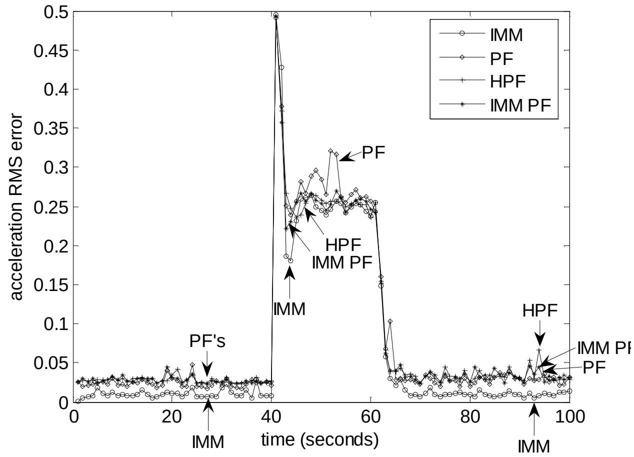


b. 10^3 particles

Figure 8. Scenario 4. The target accelerates with 0.1 g after 40 s. The particle filter parameters are $\sigma_a = 0.1g$, $\tau_1 = 5000\text{s}$ and $\tau_2 = 500\text{s}$.

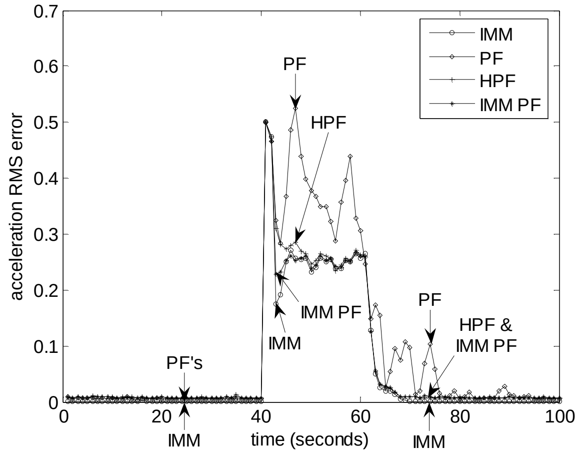


a. 10^4 particles

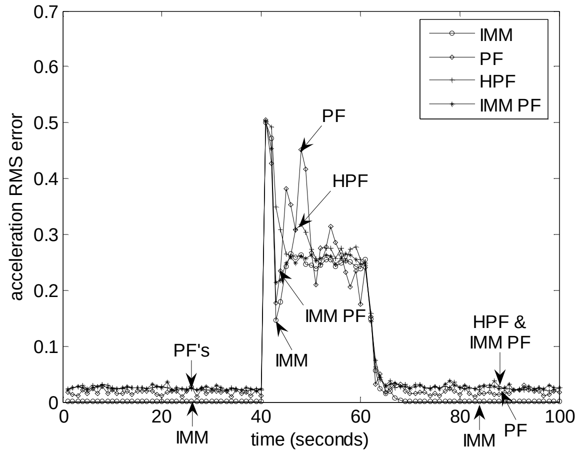


b. 10^3 particles

Figure 9. Scenario 1. The target accelerates with 5g between 40 s and 60 s. The particle filter parameters are $\sigma_a = 5\text{g}$, $\tau_1 = 50\text{s}$ and $\tau_2 = 5\text{s}$.

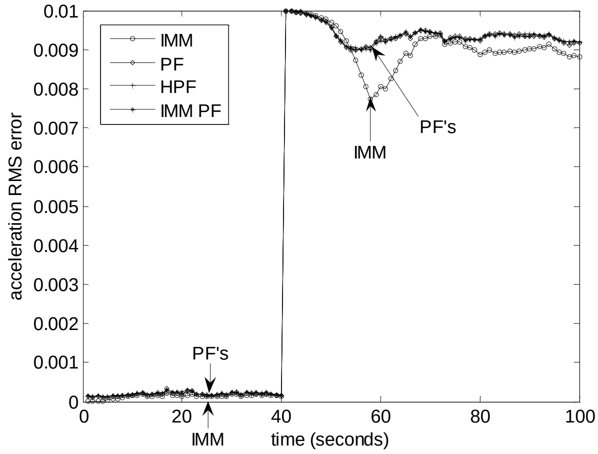


a. 10^4 particles

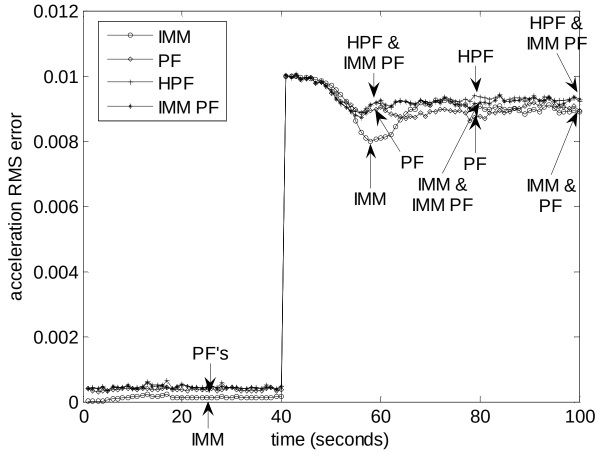


b. 10^3 particles

Figure 10. Scenario 2. The target accelerates with 5g between 40 s and 60 s. The particle filter parameters are $\sigma_a = 5\text{g}$, $\tau_1 = 5000\text{s}$ and $\tau_2 = 5\text{s}$.

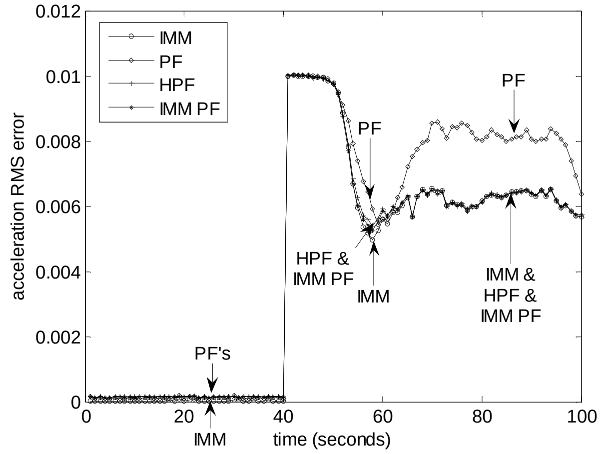


a. 10^4 particles

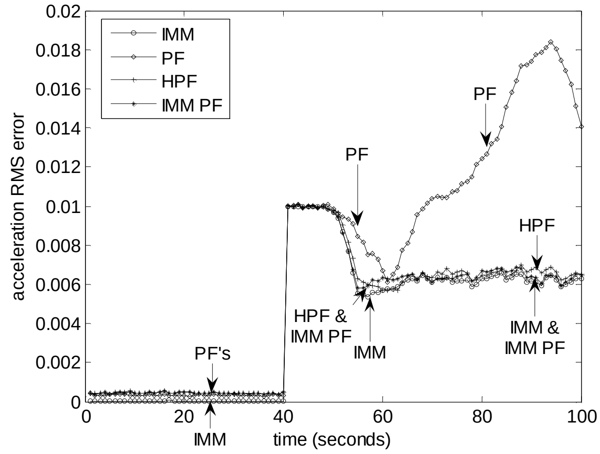


b. 10^3 particles

Figure 11. Scenario 3. The target accelerates with 0.1 g after 40 s. The particle filter parameters are $\sigma_a = 0.1\text{g}$, $\tau_1 = 50\text{s}$ and $\tau_2 = 5\text{s}$.



a. 10^4 particles



b. 10^3 particles

Figure 12. Scenario 4. The target accelerates with 0.1 g after 40 s. The particle filter parameters are $\sigma_a = 0.1\text{g}$, $\tau_1 = 5000\text{s}$ and $\tau_2 = 500\text{s}$.