State Estimation for Stochastic Hybrid Systems Based on Deterministic Dirac Mixture Approximation

Maxim Dolgov[1], Gerhard Kurz[1], and Uwe D. Hanebeck[1]

***Abstract*— In this paper, we consider state estimation for Stochastic Hybrid Systems (SHS). These are systems that possess both continuous-valued and discrete-valued dynamics. For SHS with nonlinear hybrid dynamics and/or non-Gaussian disturbances, state estimation can be implemented as an Interacting Multiple Model (IMM) particle filter. However, a disadvantage of particle filtering is the computational load caused by the large number of particles required for a sufficiently good estimation. We address this issue by first expressing the probability density that describes the state of the SHS as a collection of densities of the continuous-valued state only conditioned on the discrete-valued state. Then, we deterministically approximate these individual densities with Dirac mixtures. The employed approximation method places the particles so that a so called modified Cramér-von Mises distance between the true and the approximated density is minimized. Deterministic approximation requires far less particles than the stochastic sampling used by particle filters. To avoid particle degeneration that can occur when a density is multiplied with the likelihood, the filter uses progressive density correction. The presented filter is demonstrated in a numerical maneuvering target tracking example.**

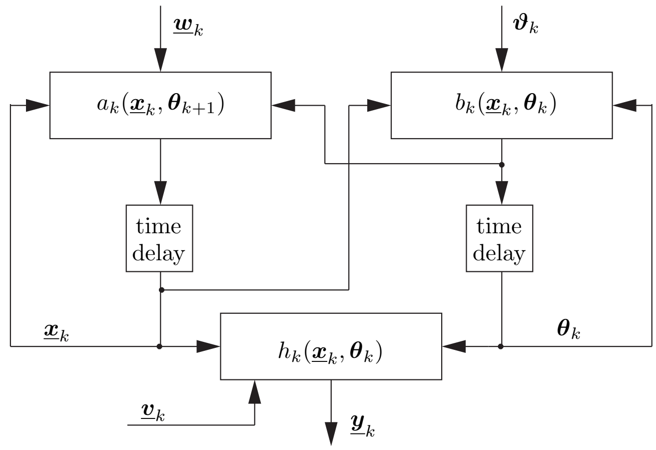

Fig. 1. Schematic of the considered Stochastic Hybrid System.

I. Introduction

Stochastic systems whose dynamics are described by continuous and discrete variables are referred to as stochastic hybrid systems (SHS). This modeling approach is often utilized when the continuous-valued state of a dynamic system not only depends on previous states and control inputs but also on a discrete-valued parameter that describes the mode of operation [1]. Examples of such systems are among others maneuvering targets, chemical processes, or switching communication networks [2]. In many practical operations the state of a hybrid system is not directly available but rather has to be estimated from measurements corrupted by noise.

In this paper, we consider a discrete-time SHS whose discrete-valued state is modeled as a hybrid Markov chain. State transitions of a hybrid Markov chain are a function of the hybrid state of the SHS. The task is to estimate both the continuous-valued state as well as the discrete-valued state based on continuous-valued measurements of the continuous-valued state only. Such a state estimation can be performed by a Multiple Model filter. However, this method is practically infeasible because the number of individual filters increases exponentially in time and therefore Interacting Multiple Model (IMM) filters are used. IMM filters compute state estimates as a weighted sum of state estimates provided by its subfilters that can exchange information among each other. Hereby, the number of subfilters is constant and finite. An overview of IMM filtering techniques is given in [3].

The filtering algorithm is usually divided in a prediction step and a filter step. In the prediction step, the filter predicts the state of the system based on the information available up to previous time step by means of the Chapman-Kolmogorov equation. However, a closed-form evaluation of this equation is usually not possible for nonlinear system dynamics and/or arbitrary system noises. An approach in this case is to approximate the occurring densities by a collection of samples (or particles) drawn from these densities [4]. By doing so the evaluation of the Chapman-Kolmogorov equation reduces to an integration of a sum of weighted Dirac impulses that represent the sampled density with the transition density that may be continuous or a sum of weighted Diracs as well. However, density approximation using Dirac mixtures can cause problems in the filter step. This issue and a possible solution will be discussed later in the introduction.

A. *Related Work*

Before IMM particle filtering (IMM-PF) was applied to SHS, it was applied to Markov Jump Systems (MJS). The discrete-valued dynamics of MJS can be modeled as a Markov chain that is, in contrast to SHS, independent from the continuous-valued dynamics. IMM-PF for MJS were, e.g., presented in [5], [6], [7]. With regard to state estimation, the fact that the continuous-valued dynamics of a MJS are independent from its discrete-valued dynamics facilitates the

[1]The authors are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, Email: maxim.dolgov@kit.edu, gerhard.kurz@kit.edu, uwe.hanebeck@ieee.org

evaluation of the Chapman-Kolmogorov equation for the continuous-valued state.

IMM particle filtering for SHS, whose discrete-valued dynamics are not independent from the continuous-valued state, was introduced in [8] and further developed in [9]. In these works, an IMM-PF for a nonlinear SHS with Gaussian noises was presented. This filter evaluates exact Bayesian equations for the discrete-valued system state. Thus, the particles are only needed for the continuous-valued state. This approach is more elegant than defining particles on the hybrid state space as, e.g., in [10], because it reduces the number of particles required for approximation.

A recent contribution in the field of IMM-PF was presented in [11]. In this work, state estimation of a continuous-time nonlinear SHS is considered. The filter employs a feedback particle filter (FPF) algorithm where the continuous-valued state is represented by particles.

A drawback of particle filtering is the large amount of samples (often more than 10^3 even for non-hybrid stochastic systems with low-dimensional state) that is necessary to provide sufficiently good estimation results. These large particle numbers are caused by the way samples representing the approximated probability density are obtained: PF use stochastic sampling to approximate a density and therefore require a large amount of samples to provide high approximation confidence. An alternative approach to stochastic sampling is the deterministic Dirac mixture density approximation that was proposed in [12], [13]. The idea of this method is to approximate a density by placing samples so that a specific value function – the modified Cramér-von Mises distance – is minimized. By making the sampling process deterministic, the number of samples can essentially be reduced while the desired approximation quality can be attained. However, the classical Cramér-von Mises distance cannot be directly applied because the cumulative distribution of multidimensional Dirac mixtures is not unambiguous. Using the so called localized cumulative distribution (LCD) that was first proposed in [14] instead of the cumulative distribution, and modifying the Cramér-von Mises distance solves this problem. Recent results in deterministic Dirac mixture approximation allow efficient approximation of Gaussian densities [15].

If a probability density is approximated with particles, stochastically or deterministically placed, it is possible that almost all particles will have negligible weights after being multiplied with the likelihood in the filter step. This phenomenon is referred to as particle degeneration. For deterministic Dirac mixtures, particle degeneration can be a greater issue because of the smaller number of particles compared to a common particle filter. As a solution to this problem, the authors in [16] proposed to employ an iterative procedure, the progressive correction. The idea of this approach is to decompose the likelihood into a product of functions. Then, at every iteration step, the current Dirac mixture is reapproximated after being multiplied with a likelihood factor. By doing so, the particles are "shifted" to the areas with high probability masses.

B. Key Idea

The filter presented in this paper estimates the hybrid state of a SHS combining the advantages of the IMM particle filtering methods and the deterministic Dirac mixture density approximation. These advantages are:

1) The probability density of the considered SHS can be decomposed in a set of continuous densities for every possible realization of the discrete-valued state. This allows to implement a filter with an IMM structure.
2) The number of samples used for approximation of the occurring densities is essentially reduced by employing deterministic Dirac mixture density approximation. To avoid particle degeneration, we substitute the multiplication with the likelihood by a progressive density correction.

Notation: To provide a consistent notation, vector-valued quantities are underlined $\underline{v}$ and random variables are in bold letters $\boldsymbol{a}$, $\underline{\boldsymbol{w}}$. A sequence $[a_k, a_{k+1}, ..., a_{k+N}]$ is abbreviated by $a_{k:k+N}$. Probability density functions of continuous-valued quantities are denoted by $f(\cdot)$ and those of the discrete-valued quantities by $p(\cdot)$.

Outline: The remainder of the paper is organized as follows. In Sec. II we define the considered system and describe the proposed solution. The relevant results on deterministic Dirac mixture approximation and progressive density correction are summarized in Sec. III. Filter equations are given in Sec. IV and the results of the numerical evaluation are presented in Sec. V.

II. Problem Formulation and Proposed Solution

A. Problem Formulation

We consider a discrete-time stochastic nonlinear system with the schematic as depicted in Fig. I that possesses a continuous-valued state $\underline{\boldsymbol{x}}_k \in \mathbb{R}^n$ and a discrete-valued state $\boldsymbol{\theta}_k \in \{1, ..., M\}, M \in \mathbb{N}$. The continuous-valued state evolves according to

$$\underline{\boldsymbol{x}}_{k+1} = a_k\left(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_{k+1}\right) + \underline{\boldsymbol{w}}_k \ , \tag{1}$$

where $\underline{\boldsymbol{w}}_k \sim f^w(\cdot)$ denotes an arbitrary additive i.i.d. noise vector that is independent of $\underline{\boldsymbol{x}}_k$ and $\boldsymbol{\theta}_k$. The evolution of $\underline{\boldsymbol{x}}_k$ described by the nonlinear function $a_k(\cdot,\cdot)$ is governed by the discrete-valued state $\boldsymbol{\theta}_k$. We will refer to $\boldsymbol{\theta}_k$ as the *mode* of the system. The mode $\boldsymbol{\theta}_k$ is assumed to be the state of a hybrid Markov chain with

$$\boldsymbol{\theta}_{k+1} = b_k(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k, \boldsymbol{\vartheta}_k) \ ,$$

where $\boldsymbol{\vartheta}_k \sim f^{\vartheta}(\cdot)$ is an independent stochastic process with $\boldsymbol{\vartheta}_k \in \mathbb{R}$. Both $\underline{\boldsymbol{x}}_k$ and $\boldsymbol{\theta}_k$ are not directly accessible but measured by a nonlinear function $h_k(\cdot,\cdot)$ according to

$$\underline{\boldsymbol{y}}_k = h_k\left(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k\right) + \underline{\boldsymbol{v}}_k \ , \tag{2}$$

where $\underline{\boldsymbol{v}}_k \in \mathbb{R}^m, \underline{\boldsymbol{v}}_k \sim f^v(\cdot)$ is an arbitrary i.i.d. additive noise that is independent of other quantities.

The goal is to recursively estimate the probability density $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k | \underline{\boldsymbol{y}}_{1:k})$ from available measurements $\underline{\boldsymbol{y}}_k$ given the initial density $f(\underline{\boldsymbol{x}}_0, \boldsymbol{\theta}_0)$.

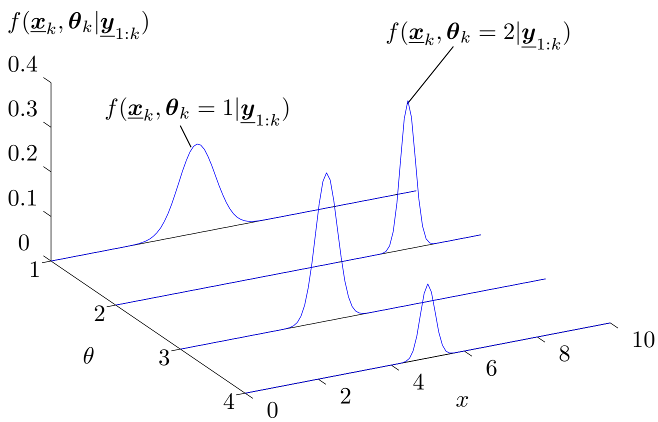

Fig. 2. The hybrid state probability density $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k | \underline{\boldsymbol{y}}_{1:k})$ defined on $\mathbb{R}^n \times \mathbb{M}$ can be interpreted as a collection of densities $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_{k-1} = j | \underline{\boldsymbol{y}}_{1:k-1})$, $j \in \{1, ..., M\}$ where each individual density is defined on $\mathbb{R}^n$ only.

B. Proposed Solution

The two main challenges in state estimation for stochastic hybrid systems are

i) the hybrid nature of the space on which $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k | \underline{\boldsymbol{y}}_{1:k})$ is defined, and

ii) the non-linearity of the equations (1) and (2).

In this paper, we address problem (i) by employing the IMM approach with M subfilters that maintain the probability densities $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = i | \underline{\boldsymbol{y}}_{1:k}), i \in \mathbb{M} = \{1, ..., M\}$. By doing so, the filter does not have to maintain a hybrid density defined on $\mathbb{R}^n \times \mathbb{M}$ but a collection of M densities each only defined on $\mathbb{R}^n$. This issue as demonstrated in Fig. 2.

To address problem (ii), we approximate the densities $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = i | \underline{\boldsymbol{y}}_{1:k})$ by means of a deterministic Dirac mixture as described in Sec. III.

III. Deterministic Dirac Mixture Approximation

In this section, we summarize relevant results on deterministic Dirac mixture density approximation and the progressive density correction. These results are only available for non-hybrid stochastic systems and will be extended to SHS in Sec. IV.

A. Deterministic Density Approximation

As mentioned in the introduction, it is possible to deterministically approximate a probability density function by a Dirac mixture. For this purpose, the approximation problem is converted into an optimization problem. The task is to place particles so that a specific approximation measure – in this case the modified Cramér-von Mises distance (CvMD) – is minimized. The unmodified CvMD evaluates the distance between the cumulative distribution of the true density and its approximation. The usage of the CvMD for approximation of a continuous density by particles is not possible, because the cumulative density of a Dirac mixture is not unambiguous. A solution to this problem is to use the so called Localized Cumulative Distance (LCD) and reformulate the CvMD as modified Cramér-von Mises distance (mCvMD).

Definition 1 *For a density function* $f(\underline{x})$, $\underline{x} \in \mathbb{R}^n$, *the* ***Localized Cumulative Distribution (LCD)*** *is defined as*

$$F(\underline{m}, \underline{b}) = \int_{\mathbb{R}^n} f(\underline{x}) \cdot K(\underline{x} - \underline{m}, \underline{b}) \, d\underline{x} \ ,$$

where $K(\cdot, \cdot)$ *is a symmetric and integrable kernel function,* $\underline{m}$ *is an offset, and* $\underline{b} \in \mathbb{R}_+^n$ *is a scaling factor.*

We restrict ourselves to an equal scaling of the kernel in each direction, thus having a scalar $b \in \mathbb{R}_+$. As a kernel, we will use a multidimensional unnormalized Gaussian given by

$$K(\underline{x} - \underline{m}, b) = \prod_{k=1}^{n} \exp\left(-\frac{1}{2} \frac{(x^{(k)} - m^{(k)})^2}{b^2} \right) \ .$$

Having defined the LCD, we can define the mCvMD.

Definition 2 *The* ***Modified Cramér-von Mises Distance (mCvMD)*** $d(F, \tilde{F})$ *of two LCDs* $F(\underline{m}, b)$ *and* $\tilde{F}(\underline{m}, b)$ *is given by*

$$d(F, \tilde{F}) = \int_{\mathbb{R}_+} \int_{\mathbb{R}^n} w(b) \left(F(\underline{m}, b) - \tilde{F}(\underline{m}, b) \right)^2 \, d\underline{m} \, db \ ,$$

where $w(b)$ *is a suitable weighting function.*

The positions $\underline{x}^\alpha$, $\alpha \in \{1, \dots, \mathcal{A}\}, \mathcal{A} \in \mathbb{N}$ of the components of the Dirac mixture

$$\widetilde{f}(\underline{\boldsymbol{x}}_k) = \sum_{\alpha=1}^{\mathcal{A}} w_k^\alpha \delta(\underline{\boldsymbol{x}}_k - \underline{x}_k^\alpha), \tag{3}$$

with

$$\sum_{\alpha=1}^{\mathcal{A}} w_k^\alpha = 1, \ w_k^\alpha > 0$$

that approximates a probability density $f(\underline{\boldsymbol{x}}_k)$ by $\mathcal{A}$ particles are the solution of the optimization problem

$$\underset{[\underline{x}^1, \dots, \underline{x}^{\mathcal{A}}]}{\arg\min} \, d(F, \tilde{F}) \ , \tag{4}$$

where F is the LCD of the density $f(\underline{\boldsymbol{x}}_k)$ and $\widetilde{F}$ is the LCD of the Dirac mixture (3) that approximates $f(\underline{\boldsymbol{x}}_k)$.

The computation of (4) can be numerically difficult. An efficient algorithm for Gaussian densities is derived in [15]. Detailed derivations of LCD and mCvMD are given in [14].

B. Progressive Density Correction

When performing Bayesian estimation, the posterior probability density is obtained by a multiplication of the prior density with the likelihood followed by normalization. If the prior probability density is given by means of a Dirac mixture, particle degeneration can occur, i.e., after the multiplication almost all particles have negligible weights. To prevent particle degeneration, we will apply progressive density correction [16]. The main idea of the progressive density correction is to decompose the likelihood $l(\underline{\boldsymbol{x}}_k) = f(\underline{\boldsymbol{y}}_k|\underline{\boldsymbol{x}}_k)$ in $s \in \mathbb{N}$ factors according to

$$l(\underline{\boldsymbol{x}}_k) = \prod_{i=1}^{s} l(\underline{\boldsymbol{x}}_k)^{\lambda_i} \ , \ \lambda_i > 0 \ \text{ and } \ \sum_{i=1}^{s} \lambda_i = 1 \ .$$

Then, instead of directly multiplying the Dirac mixture with the likelihood $l(\underline{\boldsymbol{x}}_k)$, s correction steps are performed. At each step, the Dirac mixture is multiplied by $l(\underline{\boldsymbol{x}}_k)^{\lambda_i}$ and reapproximated. By doing so, the particles are "shifted" to the areas with high posterior probability mass.

A detailed description of the algorithm and of the method how correction exponents λ_i are calculated can be found in [16] and [17].

IV. The SHS Filter Equations

In this section, we derive the filter equations for the considered SHS. First, the equations are given for arbitrary probability densities of the continuous-valued state. However, in this case, the equations cannot be evaluated in closed form. Therefore, we approximate the state and the system noise densities by a deterministic Dirac mixture. Finally, we summarize the estimation steps that are performed at each time step.

A. Filter Equations for Arbitrary Densities

Assume a collection of prior probability densities $f(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i|\underline{\boldsymbol{y}}_{1:k-1})$, $i \in \{1, \ldots, M\}$ is given. Then, a filtered density $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j|\underline{\boldsymbol{y}}_{1:k})$ can be computed by

$$\begin{aligned} f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j|\underline{\boldsymbol{y}}_{1:k}) \\ = f_L(\underline{\boldsymbol{y}}_{1:k}, \underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j) f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j|\underline{\boldsymbol{y}}_{1:k-1}) \end{aligned} \quad (5)$$

with the normalized likelihood

$$f_L(\underline{\boldsymbol{y}}_{1:k}, \underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j) = \frac{f(\underline{\boldsymbol{y}}_k|\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j)}{f(\underline{\boldsymbol{y}}_k|\underline{\boldsymbol{y}}_{1:k-1})}$$

and the prediction

$$\begin{aligned} & f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j|\underline{\boldsymbol{y}}_{1:k-1}) \\ & = \int_{\mathbb{R}^n} \sum_{i=1}^{M} f(\underline{\boldsymbol{x}}_k|\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_k = j) p(\boldsymbol{\theta}_k = j|\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i) \\ & \quad \times f(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i|\underline{\boldsymbol{y}}_{1:k-1}) \, \mathrm{d}\underline{\boldsymbol{x}}_{k-1} \ . \end{aligned} \quad (6)$$

The derivation of this equations is described in the appendix. The unnormalized likelihood $f(\underline{\boldsymbol{y}}_k|\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j)$, and the transition probabilities $f(\underline{\boldsymbol{x}}_k|\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_k = j)$ and $p(\boldsymbol{\theta}_k = j|\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i)$ are given by

$$\begin{aligned} & f(\underline{\boldsymbol{y}}_k|\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j) = f^v(\underline{\boldsymbol{y}}_k - h_k(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j)) \ , \\ & f(\underline{\boldsymbol{x}}_k|\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_k = j) = f^w(\underline{\boldsymbol{x}}_k - a_k(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_k = j)) \ , \\ & p(\boldsymbol{\theta}_k = j|\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i) \\ & = \int_{\mathbb{R}} \mathbb{1}(\boldsymbol{\theta}_k = j, b_k(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i, \boldsymbol{u}_{k-1})) f^{\vartheta}(\boldsymbol{\vartheta}_k) \, \mathrm{d}\boldsymbol{\vartheta}_k \ , \end{aligned}$$

with the indicator function

$$\mathbb{1}(m, n) = \begin{cases} 1, & m = n \ , \\ 0, & \text{otherwise} \ . \end{cases}$$

The calculation of these transition densities is described in the appendix.

B. Filter Equations for Approximated Densities

The evaluation of (6) cannot be performed in closed form for arbitrary probability density function. We therefore approximate the density of the hybrid state $f(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i|\underline{\boldsymbol{y}}_{1:k-1})$ and the densities of the system noise $f^w(\underline{\boldsymbol{w}}_{k-1})$ by means of a deterministic Dirac mixture (3). The densities are then given by

$$\widetilde{f}(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i|\underline{\boldsymbol{y}}_{1:k-1}) = \sum_{\alpha=1}^{\mathcal{A}} w_{k-1}^{i,\alpha} \delta(\underline{\boldsymbol{x}}_{k-1} - \underline{x}_{k-1}^{i,\alpha}) \quad (7)$$

and

$$\widetilde{f}^w(\underline{\boldsymbol{w}}_{k-1}) = \sum_{\beta=1}^{\mathcal{B}} w_{k-1}^{\beta} \delta(\underline{\boldsymbol{w}}_{k-1} - \underline{x}_{k-1}^{\beta}) \ , \quad (8)$$

where $w_{k-1}^{\alpha}, w_{k-1}^{\beta}$ denote the weights of the components of the Dirac mixture and $\underline{x}_{k-1}^{\alpha}, \underline{x}_{k-1}^{\beta}$ their positions. The numbers of components are $\mathcal{A} \in \mathbb{N}$ and $\mathcal{B} \in \mathbb{N}$, respectively.

With (7) and (8) the prediction step in (6) becomes

$$\begin{aligned} & \widetilde{f}(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j|\underline{\boldsymbol{y}}_{1:k-1}) \\ & = \sum_{i=1}^{M} \sum_{\alpha=1}^{\mathcal{A}} \sum_{\beta=1}^{\mathcal{B}} w_{k-1}^{i,\alpha} w_{k-1}^{\beta} \delta(\underline{\boldsymbol{x}}_k - a_k(\underline{x}_{k-1}^{i,\alpha}, \boldsymbol{\theta}_k = j) - \underline{x}_{k-1}^{\beta}) \\ & \times p(\boldsymbol{\theta}_k = j|\underline{x}_{k-1}^{i,\alpha}, \boldsymbol{\theta}_{k-1} = i) \ . \end{aligned}$$

The approximation of the state densities by a Dirac mixture not only allows for processing of arbitrary state and system noise densities but also facilitates the calculation of the transition density $p(\boldsymbol{\theta}_k = j|\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i)$ because the indicator function $\mathbb{1}(\cdot, \cdot)$ only has to be evaluated at discrete points $(\underline{x}_{k-1}^{\alpha}, \boldsymbol{\theta}_{k-1} = i)$.

As mentioned in the introduction, multiplying the predicted density given by means of a Dirac mixture can lead to particle degeneration. To address this issue, we propose to implement the filter step as a progressive correction as proposed in [16].

C. Implementation

At each time step, the presented filter performs four steps to estimate the probability densities $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j|\underline{\boldsymbol{y}}_{1:k})$ based on the previously estimated densities $f(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i|\underline{\boldsymbol{y}}_{1:k-1})$ and the measurement $\underline{\boldsymbol{y}}_k$. Steps 1 to 3 constitute the predictions step while step 4 constitutes the filter step.

Step 1: For each mode $\boldsymbol{\theta}_k = j$, propagate the samples $\delta(\underline{\boldsymbol{x}}_k - \underline{x}^{i,\alpha})$ through the system equation using the information from each previous mode $\boldsymbol{\theta}_{k-1} = i$

$$\sum_{i=1}^{M} \sum_{\alpha=1}^{\mathcal{A}} w_{k-1}^{i,\alpha} \delta(\underline{\boldsymbol{x}}_{k-1} - \underline{x}_{k-1}^{i,\alpha}) \longrightarrow \sum_{i=1}^{M} \sum_{\alpha=1}^{\mathcal{A}} w_{k-1}^{i,\alpha} \delta(\underline{\boldsymbol{x}}_k - a_k(\underline{x}_{k-1}^{i,\alpha}, j)) \ .$$

Step 2: Convolve the results from Step 1 with the Dirac mixture of the system noise

$$\sum_{i=1}^{M}\sum_{\alpha=1}^{\mathcal{A}} w_{k-1}^{i,\alpha}\delta(\underline{\boldsymbol{x}}_k - a_k(\underline{x}_{k-1}^{i,\alpha}, j)) \longrightarrow$$
$$\sum_{i=1}^{M}\sum_{\alpha=1}^{\mathcal{A}}\sum_{\beta=1}^{\mathcal{B}} w_{k-1}^{i,\alpha} w_{k-1}^{\beta}\delta(\underline{\boldsymbol{x}}_k - a_k(\underline{x}_{k-1}^{i,\alpha}, \boldsymbol{\theta}_k = j) - \underline{x}_{k-1}^{\beta}) .$$

Step 3: Finalize the prediction step by computing the influence of the hybrid Markov chain on the weights of the samples.

$$\sum_{i=1}^{M}\sum_{\alpha=1}^{\mathcal{A}}\sum_{\beta=1}^{\mathcal{B}} w_{k-1}^{i,\alpha} w_{k-1}^{\beta}\delta(\underline{\boldsymbol{x}}_k - a_k(\underline{x}_{k-1}^{i,\alpha}, \boldsymbol{\theta}_k = j) - \underline{x}_{k-1}^{\beta})$$
$$\longrightarrow \sum_{i=1}^{M}\sum_{\alpha=1}^{\mathcal{A}}\sum_{\beta=1}^{\mathcal{B}} w_{k-1}^{i,\alpha} w_{k-1}^{\beta}\delta(\underline{\boldsymbol{x}}_k - a_k(\underline{x}_{k-1}^{i,\alpha}, \boldsymbol{\theta}_k = j) - \underline{x}_{k-1}^{\beta})$$
$$\times\, p(\boldsymbol{\theta}_k = j | \underline{x}_{k-1}^{i,\alpha}, \boldsymbol{\theta}_{k-1} = i) \ .$$

Step 4: Perform progressive correction according to Alg. 1 given in [16].

V. Evaluation

In order to evaluate the filter, a Monte-Carlo simulation with 4000 runs à 40 time steps each was performed. As a reference an IMM particle filter was used that has the same IMM structure as the proposed filter but approximates the individual densities $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j | \underline{\boldsymbol{y}}_{1:k})$ using common methods for stochastic sampling. The equations (1) and (2) of the considered SHS were chosen to

$$\boldsymbol{x}_{k+1} = \begin{cases} 0.9{\cdot}\boldsymbol{x}_k - 0.5\sqrt{|\boldsymbol{x}_k|} + \boldsymbol{w}_k, & \text{if } \boldsymbol{\theta}_{k+1} = 1 \\ 0.9{\cdot}\boldsymbol{x}_k + 0.5\sqrt{|\boldsymbol{x}_k|} + \boldsymbol{w}_k, & \text{if } \boldsymbol{\theta}_{k+1} = 2 \\ 0.9{\cdot}\boldsymbol{x}_k + \boldsymbol{w}_k, & \text{if } \boldsymbol{\theta}_{k+1} = 3 \end{cases} ,$$
$$\boldsymbol{y}_k = \text{sgn}(\boldsymbol{x}_k)\sqrt{|\boldsymbol{x}_k|} + \boldsymbol{v}_k \ ,$$

with zero-mean Gaussian noise terms $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$.

The evolution of the discrete-valued state was governed by a classical Markov chain with the transition matrix

$$T = \begin{bmatrix} 0.4 & 0.2 & 0.4 \\ 0.1 & 0.3 & 0.6 \\ 0.2 & 0.1 & 0.7 \end{bmatrix} ,$$

with

$$T_{ij} = \text{Prob}(\boldsymbol{\theta}_k = j | \boldsymbol{\theta}_{k-1} = i) \ .$$

For each simulation run, both filters were initialized with

$$\text{E}\{\boldsymbol{x}_0\} = 0 \text{ and } \text{E}\left\{(\boldsymbol{x}_0 - \text{E}\{\boldsymbol{x}_0\})(\boldsymbol{x}_0 - \text{E}\{\boldsymbol{x}_0\})^T\right\} = 5^2,$$
$$\text{Prob}(\boldsymbol{\theta}_0 = 1) = 0,\ \text{Prob}(\boldsymbol{\theta}_0 = 2) = 1,\ \text{Prob}(\boldsymbol{\theta}_0 = 3) = 0 \ .$$

In simulations, we analyzed (i) the influence of the variation of the number of particles used for density approximation as well as (ii) the influence of the system and the measurement noise on the estimation quality. To analyze (i), we tested the proposed filter with $\mathcal{A} = 5$ particles per continuous-valued density $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = i | \underline{\boldsymbol{y}}_{1:k})$ and $\mathcal{B} = 5$ particles for the system noise density $f^w(\underline{\boldsymbol{w}}_k)$ as well as $\mathcal{A} = 10$ and $\mathcal{B} = 5$. The IMM particle filter on the other hand used 100 and 1000 particles respectively. For the analysis of (ii), different values for the standard deviations of the system and the measurement noises were used.

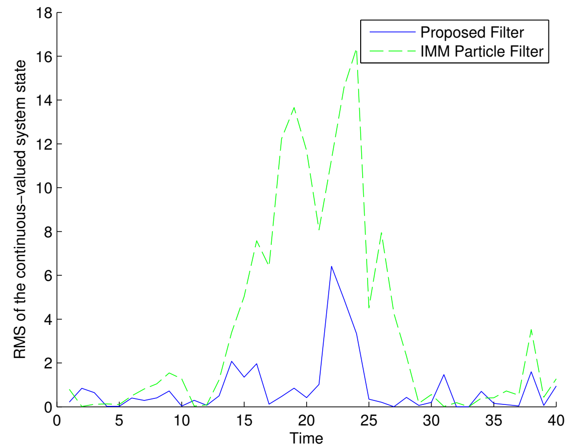

Fig. 3. Depicted is the quadratic error of the continuous-valued state estimate. The system noise is $\sigma_w = 1.5$ and the measurement noise $\sigma_v = 1$. The proposed filter approximates continuous densities with 10 samples and the IMM particle filter with 10^3 samples.

The results of a representative simulation run (RMSE of the continuous-valued state) with standard deviation of the system noise $\sigma_w = 1.5$ and the standard deviation of the measurement noise $\sigma_v = 1$ are depicted in Fig. 3. For this simulation run, the particle filter used 10^3 particles per continuous-state density $f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k | \underline{\boldsymbol{y}}_{1:k})$ and the proposed filter used 10 particles per continuous state density and 5 particles for the system noise density. Table V shows the Root Mean Square (RMS) error of the continuous-valued state for all simulated scenarios. The proposed filter provides comparable results as the particle filter while requiring far less particles. In scenarios with low measurement noise, the presented filter outperforms the particle filter.

VI. Conclusion

In this paper, we presented an IMM filter that estimates the hybrid state of a stochastic hybrid system from measurements of the continuous-valued state. The considered SHS possesses arbitrary continuous-valued (linear or nonlinear) dynamics perturbed by an arbitrary additive noise. The discrete-valued state of the SHS, referred to as mode, is governed by a hybrid Markov chain. The probability density of the hybrid system state is expressed as a collection of densities of the continuous-valued state for each possible realization of the discrete-valued state. These continuous-valued densities are approximated by a deterministic Dirac mixture. To mitigate particle degeneration, the filter step is implemented as a progressive procedure.

In the simulated scenario, the proposed filter had a comparable performance as the IMM-PF while requiring much less particles for density approximation.

Noise parameters	proposed filter $\mathcal{A}=5,\ \mathcal{B}=5$	proposed filter $\mathcal{A}=10,\ \mathcal{B}=5$	IMM particle filter 100 particles	IMM particle filter 1000 particles
$\sigma_w=1,\ \sigma_v=1$	1.6265	1.6166	1.6383	1.6117
$\sigma_w=1.5,\ \sigma_v=1$	2.1460	2.0773	2.1705	2.0640
$\sigma_w=1,\ \sigma_v=1.5$	2.0752	2.0712	2.0088	1.9932

TABLE I

SIMULATION RESULTS.

APPENDIX

A. Derivation of the Filter Equation

Derivation of (5) can be performed according to

$$\begin{aligned}
&f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j | \underline{\boldsymbol{y}}_{1:k}) \\
&= \int_{\mathbb{R}^n} \sum_{i=1}^{M} f(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_k = j, \boldsymbol{\theta}_{k-1} = i | \underline{\boldsymbol{y}}_{1:k}) \ \mathrm{d}\underline{\boldsymbol{x}}_{k-1} \\
&= \frac{f(\underline{\boldsymbol{y}}_k | \underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j)}{f(\underline{\boldsymbol{y}}_k | \underline{\boldsymbol{y}}_{1:k-1})} \int_{\mathbb{R}^n} \sum_{i=1}^{M} f(\underline{\boldsymbol{x}}_k | \underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_k = j) \\
&\times p(\boldsymbol{\theta}_k{=}j | \underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1}{=}i) f(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1}{=}i | \underline{\boldsymbol{y}}_{1:k-1}) \ \mathrm{d}\underline{\boldsymbol{x}}_{k-1} \\
&= \underbrace{\frac{f(\underline{\boldsymbol{y}}_k | \underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j)}{f(\underline{\boldsymbol{y}}_k | \underline{\boldsymbol{y}}_{1:k-1})}}_{\text{filtering}} \cdot \underbrace{f(\underline{\boldsymbol{x}}_k, \boldsymbol{\theta}_k = j | \underline{\boldsymbol{y}}_{1:k-1})}_{\text{prediction}} \ .
\end{aligned}$$

B. Transition Densities of Continuous-Valued States

For a system equation of the form (1) with additive arbitrary noise $\underline{w}_{k-1}$ that is independent from other quantities, the transition probability $f(\underline{x}_k | \underline{x}_{k-1}, \boldsymbol{\theta}_k = j)$ can be calculated according to

$$\begin{aligned}
&f(\underline{x}_k | \underline{x}_{k-1}, \boldsymbol{\theta}_k = j) \\
&= \int_{\mathbb{R}^n} f(\underline{x}_k | \underline{x}_{k-1}, \boldsymbol{\theta}_k = j, \underline{w}_{k-1}) f^w(\underline{w}_{k-1}) \ \mathrm{d}\underline{w}_{k-1} \\
&= \int_{\mathbb{R}^n} \delta(\underline{x}_k - a_k(\underline{x}_{k-1}, \boldsymbol{\theta}_k = j) - \underline{w}_{k-1}) f^w(\underline{w}_{k-1}) \ \mathrm{d}\underline{w}_{k-1} \\
&= f^w(\underline{x}_k - a_k(\underline{x}_{k-1}, \boldsymbol{\theta}_k = j)) \ .
\end{aligned}$$

The calculation of the likelihood $f(\underline{y}_k | \underline{x}_k, \boldsymbol{\theta}_k = j)$ can be performed analogously.

C. Transition Densities of Discrete-Valued States

The transition probability $p(\boldsymbol{\theta}_k = j | \underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i)$ is given by

$$\begin{aligned}
&p(\boldsymbol{\theta}_k = j | \underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i) \\
&= \mathrm{E}\left\{ \mathbb{1}(\boldsymbol{\theta}_k, j) | \, \underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i \right\} \\
&= \int_{\mathbb{R}} \mathbb{1}(\boldsymbol{\theta}_k = j, b_k(\underline{\boldsymbol{x}}_{k-1}, \boldsymbol{\theta}_{k-1} = i, \boldsymbol{u}_{k-1})) f^{\vartheta}(\boldsymbol{\vartheta}_k) \ \mathrm{d}\boldsymbol{\vartheta}_k \ .
\end{aligned}$$

ACKNOWLEDGMENT

This work was supported by the German Science Foundation (DFG) within the Research Training Groups RTG 1194 “Self-organizing Sensor-Actuator-Networks” and RTG 1126 “Soft-tissue Surgery: New Computer-based Methods for the Future Workplace”.

REFERENCES

[1] P. J. Antsaklis, “A Brief Introduction to the Theory and Applications of Hybrid Systems,” in *Proceedings of the IEEE*, vol. 88, no. 7, 2000, pp. 879–887.

[2] H. A. H. Blom, Ed., *Stochastic Hybrid Systems: Theory and Safety Critical Applications*, ser. Lecture Notes in Control and Information Sciences ; 337. Berlin: Springer, 2006.

[3] X. Rong-Li and V. Jilkov, “Survey of Maneuvering Target Tracking. Part V: Multiple-model Methods,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1255–1321, 2005.

[4] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A Tutorial on Particle Filters for Online Nonlinear/non-Gaussian Bayesian Tracking,” *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[5] C. Andrieu, M. Davy, and A. Doucet, “Efficient Particle Filtering for Jump Markov Systems. Application to Time-Varying Autoregressions,” *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1762–1770, 2003.

[6] H. Driessen and Y. Boers, “An Efficient Particle Filter for Jump Markov Nonlinear Systems,” in *IEE Target Tracking 2004: Algorithms and Applications*, 2004, pp. 19–22.

[7] M. Morelande and S. Challa, “Manoeuvring Target Tracking in Clutter using Particle Filters,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 1, pp. 252–270, 2005.

[8] H. A. P. Blom and E. Bloem, “Particle Filtering for Stochastic Hybrid Systems,” in *Proceedings of the 43rd IEEE Conference on Decision and Control (CDC 2004)*, vol. 3, 2004, pp. 3221–3226.

[9] ——, “Exact Bayesian and Particle Filtering of Stochastic Hybrid Systems,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 55–70, 2007.

[10] S. McGinnity and G. Irwin, “Multiple Model Bootstrap Filter for Maneuvering Target Tracking,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 3, pp. 1006–1012, Jul 2000.

[11] T. Yang, H. A. Blom, and P. G. Mehta, “Interacting Multiple Model-Feedback Particle Filter for Stochastic Hybrid Systems,” in *Proceedings of the 52nd Conference on Decision and Control (CDC 2013)*, Dec 2013, pp. 7065–7070.

[12] O. C. Schrempf, D. Brunn, and U. D. Hanebeck, “Dirac Mixture Density Approximation Based on Minimization of the Weighted Cramer-von Mises Distance,” in *Proceedings of the 2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006)*, Heidelberg, Germany, Sep. 2006, pp. 512–517.

[13] ——, “Density Approximation Based on Dirac Mixtures with Regard to Nonlinear Estimation and Filtering,” in *Proceedings of the 2006 IEEE Conference on Decision and Control (CDC 2006)*, San Diego, California, USA, Dec. 2006.

[14] U. D. Hanebeck and V. Klumpp, “Localized Cumulative Distributions and a Multivariate Generalization of the Cramér-von Mises Distance,” in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008, pp. 33–39.

[15] I. Gilitschenski and U. D. Hanebeck, “Efficient Deterministic Dirac Mixture Approximation,” in *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington D. C., USA, Jun. 2013.

[16] P. Ruoff, P. Krauthausen, and U. D. Hanebeck, “Progressive Correction for Deterministic Dirac Mixture Approximations,” in *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, Chicago, Illinois, USA, 2011.

[17] U. D. Hanebeck, “PGF 42: Progressive Gaussian Filtering with a Twist,” in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.