Ensemble Gaussian Mixture Filter based on Projected Cramér-von Mises Distance

Uwe D. Hanebeck*, Dominik Prossel*, Andrey A. Popov†, Felipe Giraldo-Grueso‡, Renato Zanetti‡

*Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
Uwe.Hanebeck@kit.edu, Dominik.Prossel@kit.edu

†Department of Infromation & Computer Sciences
University of Hawai'i at Manoa, Honolulu, HI 96822, USA
apopov@hawaii.edu

‡Department of Aerospace Engineering and Engineering Mechanics
The University of Texas at Austin
Austin, TX 78712, USA
fgiraldo@utexas.edu, renato@utexas.edu

***Abstract*—This paper proposes modifications of the Ensemble Gaussian Mixture Filter (EnGMF). The EnGMF is a combination of a particle filter and a Gaussian sum filter. It switches between two density representations: Samples or particles are used for the prediction step (time update) while Gaussian Mixtures (GMs) are used for the filter step (measurement update). The challenge is the continual conversion between GMs and samples. The quality loss during conversion should be minimal, while conversion should also be fast. Here, we discuss a systematic deterministic sampling technique for converting a GM into a set of particles. It is based on Projected Cumulative Distributions (PCDs) that are compared with a Cramér-von Mises distance. This allows a direct approximation of GMs without intermediate sampling and an adjustable tradeoff between quality and computational complexity.**

I. Introduction

This paper is in the area of state estimation for nonlinear discrete-time dynamic systems, where we consider recursively estimating the state based on sequentially arriving observations by an appropriate filtering mechanism. Depending on the complexity of the system (degree of nonlinearity and noise structure), various filtering techniques are available in literature. These range from Kalman filter-like techniques using a Gaussian density representation for (almost) linear systems to particle filters using a sample representation for strongly nonlinear systems.

In this paper, we consider the often encountered type of systems with a strongly nonlinear system equation describing the evolution of the state over time and a weakly nonlinear measurement equation relating the state and the observations. We will argue that in this case, a single density representation is not well suited for both the prediction step (using the system model) and the filtering step (using the measurement model). Instead, we use a hybrid density representation where samples are used in the prediction step as in the particle filter and GMs are used in the filter step. The reasons for doing so are as follows (with more details in the paper): (i) Propagating uncertainty through the system model is straightforward for particle representations, while particle degeneracy in the particle filter step for including the measurement information is rather difficult to avoid. (ii) The filter step for a GM representation can be implemented efficiently, while propagating GMs through

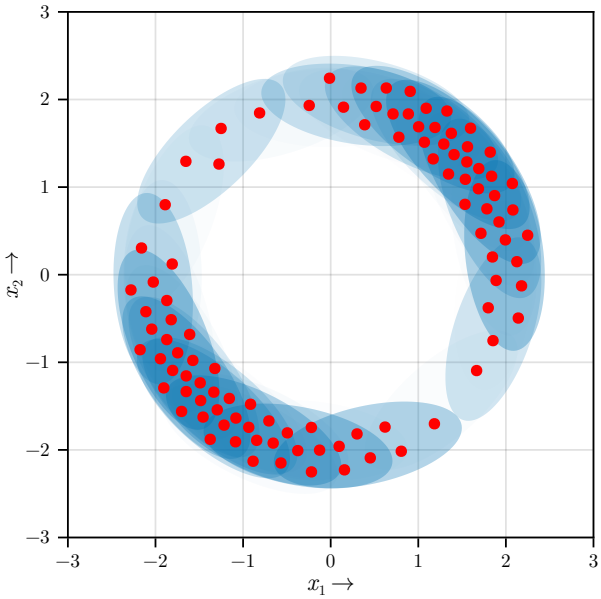

Fig. 1: Sample approximation (in red) of a Gaussian Mixture (GM) density (in blue) based on the proposed deterministic sampling technique using 100 samples.

the system model is complex, unless linearization is employed, which limits the estimation quality.

When the GM components are processed individually in the filter step by means of local filters such as the Extended Kalman Filter (EKF) or the Unscented Kalman Filter (UKF), the resulting filter structure is called the Ensemble Gaussian Mixture Filter (EnGMF) [1, 22, 17].

An obvious challenge for the EnGMF is the need for conversion between the two representations: (i) The first conversion is from particles to GMs after prediction. Kernel density estimation (KDE) methods can be used here, see [20] and its use in [22, p. 4184]. Alternatively, in the additive noise case, convolution of the noise density with the propagated samples can be used as the desired mixture representation. (ii) The second conversion requires sampling from the GM resulting from the filter step. This is significantly more complicated than the first conversion when high-quality results are desired and various options have been proposed in literature.

The simplest option would be to randomly sample from the GM, resulting in inhomogeneous samples with poor coverage. High-quality deterministic sampling based on systematically minimizing a distance measure between the given GM and the particle set has been proposed in [5]. However, it is rather computationally complex. In order to speed up the sampling, two-step techniques have been proposed [6, 7] that first sample from the individual GM components and then perform a sample reduction of the union of samples. In [6], deterministic Fibonacci grids [4] are used for sampling from the individual GM components and reduction is performed with the technique from [9]. [7] uses a transport-based reduction technique. These two-step procedures share some disadvantages: Computational complexity is still high and the intermediate sampling from individual components introduces unwanted artifacts (plus, the number of intermediate samples is another parameter that has to be fixed).

In this paper, we propose to directly sample from the GM based on Projected Cumulative Distributions (PCDs) [8]. The key idea is to simplify the distance measure between the given GM and the particle set by considering a finite number of one-dimensional density projections. The cumulative distribution functions of these projections are compared, leading to a type of Cramér-von Mises distance that can be efficiently computed and optimized. An example is shown in Fig. 1.

Besides introducing the new sampling method, the contributions of this paper include (i) a detailed derivation of the EnGMF, (ii) arguments for its usefulness in estimating the state in certain classes of dynamic systems, and (iii) a thorough comparison between several techniques for sampling from the posterior GM.

The paper is structured as follows. In the next section, we characterize the type of nonlinear dynamic system we assume for filtering. We then derive a general form of the EnGMF in Sec. III along with some simpler variants. Techniques for converting GMs to sample sets will be discussed in Sec. IV. Numerical evaluations for comparing the proposed techniques to the state-of-the-art will be conducted in Sec. V.

II. Problem Formulation

In this paper, we develop a filter method for estimating the state of nonlinear discrete-time dynamic systems from observations. Our specific focus is on systems with strong nonlinearities and potentially non-Gaussian noise in the system model. However, the measurement equation relating states and measurements is assumed to be weakly nonlinear, i.e., easy to linearize, and additive Gaussian measurement noise is preferred.

We consider the following discrete-time nonlinear dynamic system

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k, \underline{\boldsymbol{w}}_k) \ , \tag{1}$$

with random state vector $\underline{\boldsymbol{x}}_k \in \mathbb{R}^N$ at time step k, known input $\underline{u}_k$, and process noise vector $\underline{\boldsymbol{w}}_k \sim f_k^w(\underline{w}_k)$. Vectors are underlined, random quantities are in boldface.

We will see that processing becomes simpler for systems with additive noise of the form

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{u}_k) + \underline{\boldsymbol{w}}_k \ . \tag{2}$$

The system state is not directly available and is observed through measurements $\underline{\boldsymbol{y}}_k$ related to the states according to the general measurement equation

$$\underline{\boldsymbol{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{v}}_k) \tag{3}$$

with measurement noise $\underline{\boldsymbol{v}}_k \sim f_k^v(\underline{v}_k)$. The filter to be developed becomes much simpler when the noise structure in the measurement equation is additive noise according to

$$\underline{\boldsymbol{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k) + \underline{\boldsymbol{v}}_k \ , \tag{4}$$

where $\underline{\boldsymbol{v}}_k$ is assumed to be Gaussian distributed measurement noise. Measurements are obtained sequentially over time and processed recursively.

In the next section, we will derive an appropriate filter for the type of system introduced. The filter will be best suited for strong system nonlinearities and weakly nonlinear measurement equations.

III. Overview EnGMF

In this section, we will briefly introduce a general version of the EnGMF. We will start the derivation from the general Bayesian filtering equations and derive specific formulations for the appropriate density representation. A block diagram of the EnGMF and its components is shown in Fig. 2.

A. General Bayesian Filtering Equations

Bayesian filtering assumes the system and measurement models to be given in probabilistic form as distributions. A transition density $f_k^{\mathcal{T}}(\underline{x}_{k+1} \,|\, \underline{x}_k)$ is used to represent the system model (1) and its uncertainty. For the special case of additive noise as in (2), the transition density is simply given by

$$f_k^{\mathcal{T}}(\underline{x}_{k+1} \,|\, \underline{x}_k) = f_k^w\left(\underline{x}_{k+1} - \underline{a}_k(\underline{x}_k, \underline{u}_k)\right) \ . \tag{5}$$

In the case of general noise structures as in (1), derivation of $f_k^{\mathcal{T}}(\underline{x}_{k+1}|\underline{x}_k)$ is far more complex.

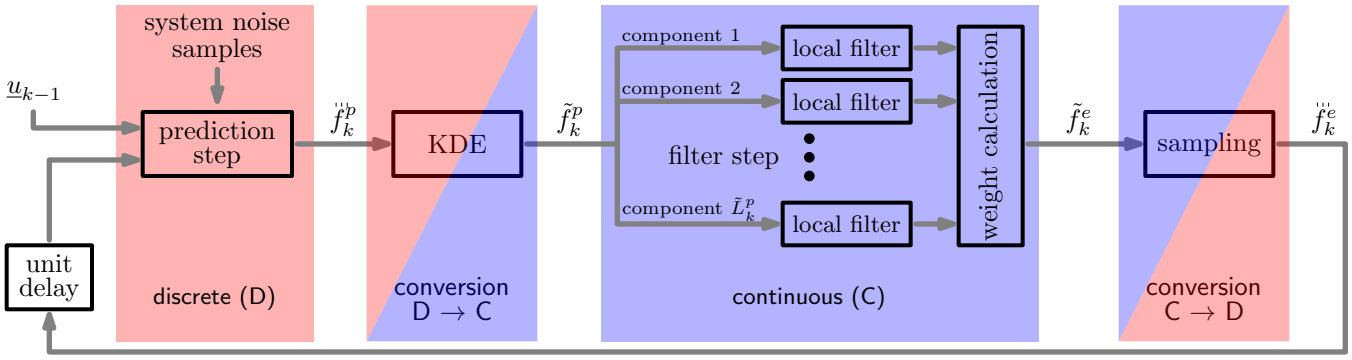

Fig. 2: Block diagram of the EnGMF and its components. It is important note that we have two different density representations for both the prediction and the filter step: A discrete particle representation and a continuous GM representation.

We define the estimated state density from the filtering step at time step k preceding the prediction step as

$$f_k^e(\underline{x}_k) = f(\underline{x}_k \,|\, \underline{y}_{1:k}, \underline{u}_{1:k}) \ , \tag{6}$$

with $\underline{y}_{1:k} = \{\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_k\}$, $\underline{u}_{1:k} = \{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_k\}$, and the prediction result from applying the system equation at time step k as

$$f_{k+1}^p(\underline{x}_{k+1}) = f(\underline{x}_{k+1} \,|\, \underline{y}_{1:k}, \underline{u}_{1:k+1}) \ . \tag{7}$$

The *general prediction step* is then given by applying the Chapman-Kolmogorov equation

$$f_{k+1}^p(\underline{x}_{k+1}) = \int_{\mathbb{R}^N} f_k^{\mathcal{T}}(\underline{x}_{k+1} \,|\, \underline{x}_k)\, f_k^e(\underline{x}_k)\, d\underline{x}_k \ . \tag{8}$$

The probabilistic form of the measurement equation (3) is the likelihood function $f_k^L(\underline{y}_k \,|\, \underline{x}_k)$. The likelihood is typically difficult to obtain from (3). The additive noise case again is the exception, and the likelihood is explicitly given by

$$f_k^L(\underline{y}_k \,|\, \underline{x}_k) = f_k^v\big(\underline{y}_k - \underline{h}_k(\underline{x}_k)\big) \ . \tag{9}$$

The general filtering step is given by Bayes' law as

$$f_k^e(\underline{x}_k) = \frac{1}{c_k}\, f_k^p(\underline{x}_k)\, f_k^L(\underline{y}_k | \underline{x}_k) \ , \tag{10}$$

where c_k is a normalization constant obtained by integration over the numerator.

B. General Thoughts

In general density function spaces for the state densities $f_k^p(\underline{x}_k)$ and $f_k^e(\underline{x}_k)$, the general Bayesian filtering equations (8) and (10) are only of theoretical value. The derivation of concrete expressions for computer implementation requires selecting appropriate density representations. Many different density representations have been proposed that all have their pros and cons. We will now discuss the major representations.

Simple parametric densities: This includes (i) the often used Gaussian representation on Euclidean domains, which is especially useful for linear or linearized systems in conjunction with linear estimators such as the Kalman filter, (ii) von Mises distributions on circular domains [14], and (iii) Bingham distributions on SO(3) [15]. For these densities, representation capabilities are typically too limited to cope with significant nonlinearities and the associated multimodality of the state densities.

Non-parametric densities: This includes mixture densities such as GMs that are known to be universal approximators, i.e., capable of approximating any density. The *prediction step* (8) is difficult to perform with mixture densities: For general transition densities, the (parameter) integral is hard to solve (especially in higher dimensions). For transition densities that are themselves GMs (with the additional difficulty that the given transition density has to be approximated by a GM), the number of components explodes and requires subsequent component reduction, unless the transition mixture components are axis-aligned [12]. However, for the *filter step* with mixture densities, various optimal [10] and suboptimal techniques for finding a posterior mixture density are available that are easy to implement.

Particles: Particles or samples can be used to represent arbitrary densities and can be viewed as a very general non-parametric representation. The employed particles are typically randomly selected, which introduces randomness into the Bayesian filtering equation that would otherwise be deterministic for a given measurement. However, deterministic particle representations have been proposed [11], which not only keep the deterministic nature of the general Bayesian filtering equations, but also have distinct advantages in terms of better coverage and higher homogeneity [3]. A deterministic version of particle resampling has been proposed in [16]. The *prediction step* (8) with particles is straightforward to perform, as every single particle can be viewed as a single realization to be propagated. At first glance, the *filter step* with particles seems to be very simple as well as it just entails multiplying the particle weights with the likelihood function at the particle locations. However, weight degeneration [21] that quickly occurs already in moderate dimensional spaces requires sophisticated update procedures [19] that are usually difficult to implement. In addition, particle-based filters typically require explicit likelihood functions, which may be difficult to obtain (for non-additive noise).

Summary: As we noticed that there is no single density representation for both the prediction and the filter step that

allows for a simple and efficient implementation in high-dimensional space, we propose to employ a hybrid representation, where we change representations between prediction and filter step. This combines the advantages of the selected representations.

C. Prediction Step (Time Update)

In the prediction step, we employ particles for representing the state density as we saw that these are well suited for strongly nonlinear forward mappings. The density of the estimate from the previous filter step (before the prediction step is performed) will be denoted by $\dddot{f}_k^e(\underline{x}_k)$ to stress that this is a discrete representation by particles. It is given by

$$\dddot{f}_k^e(\underline{x}_k) = \sum_{i=1}^{\dddot{L}_k^e} \dddot{w}_{k,i}^e \, \delta(\underline{x}_k - \dddot{\underline{x}}_{k,i}^e) \tag{11}$$

with $\delta(.)$ the Dirac delta distribution. The density is comprised of $\dddot{L}_k^e$ particles at locations $\dddot{\underline{x}}_{k,i}^e$ with weights $\dddot{w}_{k,i}^e$. Thanks to the sifting property of the Dirac delta distribution, plugging $\dddot{f}_k^e(\underline{x}_k)$ into (8) removes the integral, and we obtain

$$\tilde{f}_{k+1}^p(\underline{x}_{k+1}) = \sum_{i=1}^{\dddot{L}_k^e} \dddot{w}_{k,i}^e \, f_k^{\mathcal{T}}(\underline{x}_{k+1} \,|\, \dddot{\underline{x}}_{k,i}^e) \ , \tag{12}$$

which is a continuous mixture (and thus denoted by $\sim$). There are two different options for using this expression for $\tilde{f}_{k+1}^p(\underline{x}_{k+1})$. The first option is to keep this representation for further processing during the filter step. The usefulness of this option becomes especially apparent for additive Gaussian system noise, i.e., $f_k^w(\underline{w}_k) = \mathrm{N}(\underline{w}_k, \underline{0}, \mathbf{C}_k^w)$ is Gaussian with zero mean and covariance matrix $\mathbf{C}_k^w$. With (5) and $\tilde{L}_k^p = \dddot{L}_k^e$, we obtain

$$\tilde{f}_{k+1}^p(\underline{x}_{k+1}) = \sum_{i=1}^{\tilde{L}_k^p} \tilde{w}_{k+1,i}^p \, N\big(\underline{x}_{k+1}, \tilde{x}_{k+1,i}^p, \mathbf{C}_k^w\big) \tag{13}$$

with GM component locations $\tilde{x}_{k+1,i}^p = \underline{a}_k(\dddot{\underline{x}}_{k,i}^e, \underline{u}_k)$ and weights $\tilde{w}_{k+1,i}^p = \dddot{w}_{k,i}^e$.

The second option is to draw samples from each component in (12), thus obtaining another Dirac mixture density (DMD). When drawing a single sample per component, the number of samples is maintained, and we have $\dddot{L}_k^p = \dddot{L}_k^e$. The final DMD resulting from the prediction step is then given by

$$\dddot{f}_{k+1}^p(\underline{x}_{k+1}) = \sum_{i=1}^{\dddot{L}_k^p} \dddot{w}_{k+1,i}^p \, \delta(\underline{x}_{k+1} - \dddot{\underline{x}}_{k+1,i}^p) \ . \tag{14}$$

Instead of sampling from (12), the new samples $\dddot{\underline{x}}_{k+1,i}^p$ can be equivalently and more intuitively directly obtained by drawing system noise samples $\underline{w}_{k,i}$, $i = 1, 2, \ldots, \dddot{L}_k^e$ and mapping the samples though the original system function from (1) as

$$\dddot{\underline{x}}_{k+1,i}^p = \underline{a}_k(\dddot{\underline{x}}_{k,i}^e, \underline{u}_k, \underline{w}_{k,i}) \tag{15}$$

for $i = 1, 2, \ldots, \dddot{L}_k^e$. Only the locations of the new samples change due to the nonlinear mapping and the noise influence. However, the weights do not change, so we have $\dddot{w}_{k+1,i}^p = \dddot{w}_{k,i}^e$.

D. Filter Step (Measurement Update)

Independent of the option selected for performing the prediction step, at time step k we assume the predicted density to be a continuous mixture, specifically a Gaussian mixture of the general form

$$\tilde{f}_k^p(\underline{x}_k) = \sum_{i=1}^{\tilde{L}_k^p} \tilde{w}_{k,i}^p \, N\big(\underline{x}_k, \tilde{\underline{x}}_{k,i}^p, \tilde{\mathbf{C}}_{k,i}^p\big) \tag{16}$$

with component weights $\tilde{w}_{k+1,i}^p$, component means $\tilde{\underline{x}}_{k+1,i}$, and component covariance matrices $\tilde{\mathbf{C}}_{k+1,i}$, for $i = 1, 2, \ldots, \tilde{L}_k^p$. (16) is either directly given by (13) or by performing Kernel Density Estimation (KDE) on the DMD (14). With KDE we have to be careful in with its limitations in higher dimensions.

We want to perform the filter step (10) with a focus on mildly nonlinear measurement equations with additive noise, see (4). When the likelihood is available, we can select from a larger variety of filters including ones providing high-quality updates such as [10]. Here, we prefer to directly work with the original measurement equation (4). In addition, we want to apply a bank of local filters (see third box from the left in Fig. 2), each taking care of a single Gaussian component of (16). This definitely provides only suboptimal results, but is simple to implement and robust, and is usually sufficient for mild nonlinearities. (Optimal results are obtained when the measurement equation is linear and the noise Gaussian.)

The updates generate new Gaussian components at locations $\tilde{\underline{x}}_{k,i}^e$ with covariance matrices $\tilde{\mathbf{C}}_{k,i}^e$ for $i = 1, 2, \ldots, \tilde{L}_k^e$ with $\tilde{L}_k^e = \tilde{L}_k^p$ that are assembled into a new GM

$$\tilde{f}_k^e(\underline{x}_k) = \sum_{i=1}^{\tilde{L}_k^e} \tilde{w}_{k,i}^e \, N\big(\underline{x}_k, \tilde{\underline{x}}_{k,i}^e, \tilde{\mathbf{C}}_{k,i}^e\big) \ . \tag{17}$$

For calculating the posterior weights $\tilde{w}_{k,i}^e$, various options of different complexity are available, see [2].

As can be seen in Fig. 2, we have to convert the continuous density $\tilde{f}_k^e(\underline{x}_k)$ resulting from the filter step into a discrete particle set representation $\dddot{f}_k^e(\underline{x}_k)$. Methods for achieving this conversion with high-quality results will be discussed in the next section.

IV. High-quality Sampling from GM

In this section, we will discuss several techniques for sampling from a given GM. First, we will discuss what we expect of the resulting samples. We will then discuss some techniques from literature. Finally, we will introduce a new technique for generating samples from GMs based on PCDs.

A. Desired Properties of Samples

The number of particles drawn from the posterior GM $\tilde{f}_k^e$ should be as small as possible as it directly influences the computational complexity of each processing step. The particles (i) should represent the $\tilde{f}_k^e$ well and (ii) should be well-spaced in terms of coverage and homogeneity, In addition, we desire (iii) a graceful degradation when the number of samples decreases, e.g., we want still want a good approximation for a very small

number of particles $\dddot{L}_k^e$ on the order of the number of GM components $\tilde{L}_k^e$.

B. State-of-the-Art

We will now take a look at the state-of-the-art of generating samples from GMs. We will start with the simplest technique, random sampling. Then, we will review an optimal GM sampling technique. Finally, we will discuss two techniques that start with samples from the individual GM components and perform a reduction to the desired number of components.

1) Random Sampling: Sampling i.i.d. from a GM is very simple to perform. In the first step, we randomly select a certain component by sampling from a categorical random variable with the GM weights as probabilities. In the second step, a sample is produced from the selected Gaussian component. As the samples are independently drawn, coverage is poor and samples may cluster together. Random sampling is very fast per sample, but a large number of samples is typically required to achieved the desired quality.

2) Deterministic Sampling based on Localized Cumulative Distribution (LCD): A method for deterministically sampling from a GM has been proposed [5]. It uses the LCDs of the GM and the particle set to define a closed-form differentiable distance measure. Minimizing this distance measure provides the desired particle set. However, this minimization is rather computationally complex. This sampling method is rather slow, but provides high-quality samples.

3) Individual Sampling and LCD Reduction: In [6], the sampling step is performed in two steps. First, each Gaussian component of the GM is sampled using deterministic Fibonacci grids [4]. The first step approximates the GM as a Dirac mixture. Note that the weights of these points may differ, as each GM may have different weights. The second step is the reduction of the first Dirac mixture approximation to the desired number of points. The Dirac mixture is reduced by minimizing the modified Cramér-von Mises distance [9] between the first approximation and equally weighted new points. The minimization is performed with respect to the location of the new points. This method is faster than the direct GM sampling from [5]. However, computational complexity depends a lot on the number of intermediate samples selected. This number is a hyper-parameter and its selection (and its effect on the final quality) is not straightforward.

4) Individual Sampling and Transport Reduction: After individually sampling each GM component, the optimal coupling matrix between the first approximation and equally weighted points is found, see also [18]. In [7], an enhanced version is developed, which iteratively changes the particle locations. These methods do not require optimization and are rather fast, when only a single iteration is used. Better quality requires more iterations and thus leads to a higher computational complexity.

C. Deterministic Sampling Based on PCD

In this paper, we propose to use the sampling method from [8]. It assembles the desired distance measure between the continuous reference density and its particle approximation from one-dimensional projections. This results in simpler expressions than in [5].

More specifically, we calculate projections of the given reference density and the particle density onto unit vectors. For every one-dimensional projection, the cumulative distribution functions for both densities (which we call Projected Cumulative Distributions) are calculated and compared with a Cramér-von Mises distance, which can be calculated analytically. Combining the one-dimensional distance measures for a finite set of suitable unit vectors gives an approximation of the desired N-dimensional distance measure. A quasi-Newton method is used to find optimal particle locations, where the particle weights are assumed to be equal. This optimization problem lends itself very well to parallelization as many calculations can be carried out independently for each projection. Out-of-the-box optimization methods such as (L-)BFGS can be used to solve this optimization problem, but do not use this problem specific structure, potentially making a custom parallel implementation faster.

The PCD-based method works for general densities, but makes most sense when the projections can be calculated analytically. In this paper, we use the PCD-based method for the special case of GM densities. In this case, the projections can be efficiently calculated in closed form.

V. Evaluation

In this section, we will provide the results of numerical evaluations of the proposed deterministic sampling technique. In Subsec. V-A, we start with an isolated investigation of approximating a given GM by a set of samples. Different techniques for sample generation will be compared in terms of quality and computational complexity. In Subsec. V-B, a single filter step based on a distance measurement will be performed, visually demonstrating the high-quality estimation results. In Subsec. V-C, recursive filtering for a nonlinear dynamic system is performed with an EnGMF and the different posterior sampling techniques.

A. Four-leaf Clover

We consider a GM with four components arranged based on the Gaussian distribution

$$\mathrm{N}\left(\begin{bmatrix}\sqrt{2}\\ \sqrt{2}\end{bmatrix}, \begin{bmatrix}1.0 & 0.5\\ 0.5 & 1.0\end{bmatrix}\right), \tag{18}$$

which was rotated around the origin in steps of 90 degrees. The resulting GM has contour lines reminiscent of a four-leaf clover, see Fig. 3.

We compare four different techniques for drawing samples from this GM (different columns in Fig. 3).

- Simple random sampling as described above ("Random").
- Random sampling with a subsequent reduction with the technique from [9] ("Rand + Red"). Here, an up-sampling rate of 2.5 was used, i.e., $2.5 \cdot N$ samples were drawn from each component.
- Generating samples of individual GM components based on the LCDs with subsequent reduction with the technique

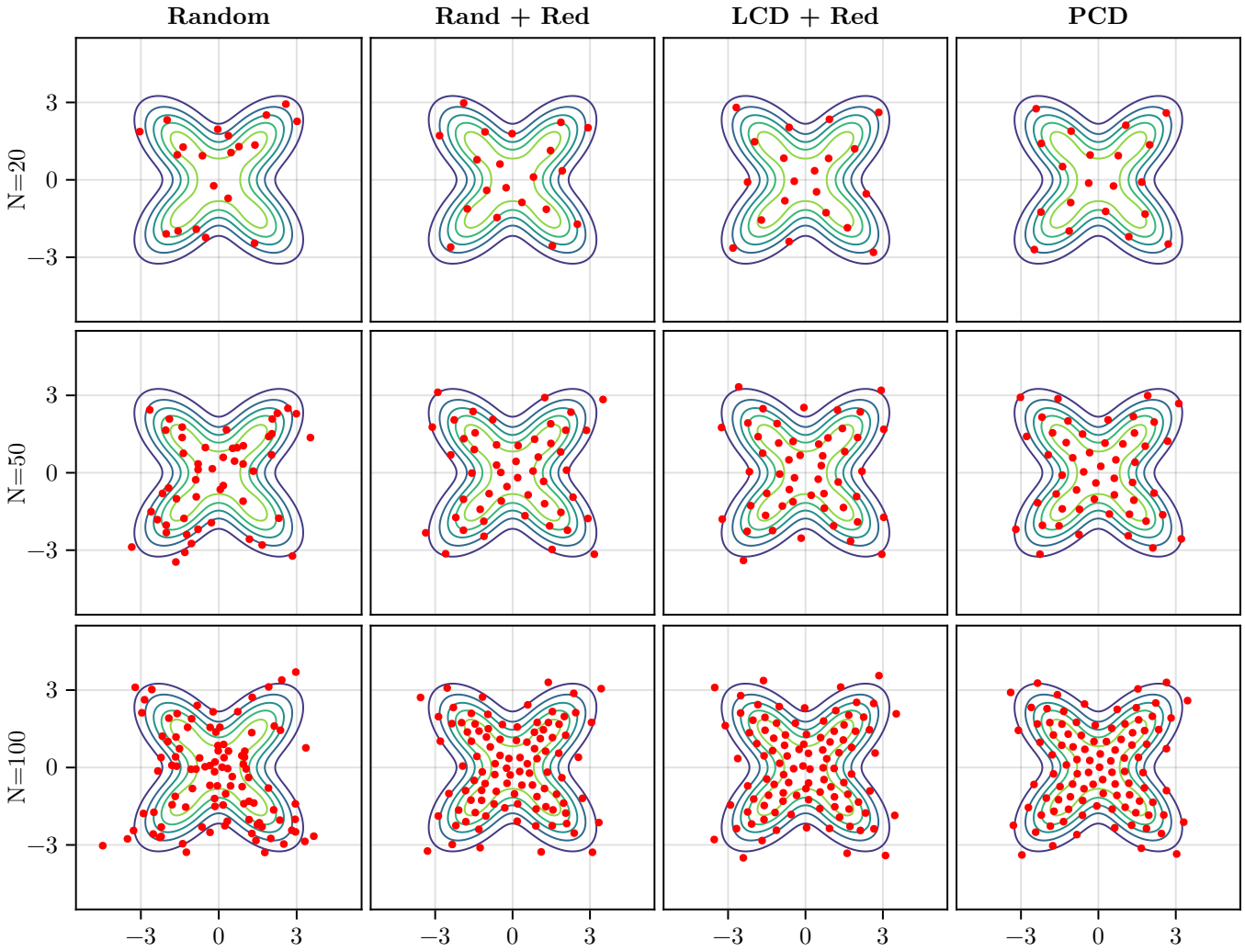

Fig. 3: Sample approximation (in red) of a GM density with different number of samples N and approximation methods.

from [9] ("LCD + Red"). Here, an upsampling rate of 0.5 was used.

- Generating samples with the proposed technique based on PCDs ("PCD"). In all cases, we generated $N = 20$, $N = 50$, and $N = 100$ samples (different rows in Fig. 3).

The results of random sampling ("Random", 1st column in Fig. 3) are typical for i.i.d. sampling and exhibit poor coverage of the given density with clusters and holes. The results of the two reduction-based two-step sampling techniques ("Rand + Red" and "LCD + Red", 2nd and 3rd column in Fig. 3) are comparable due to the larger number of random samples used in the first step. In both cases, coverage is good but homogeneity can be improved. The results of the proposed sampling technique ("PCD", 4th column in Fig. 3) are the most homogeneous and cover the given density very well.

B. Double Banana Shape Induced by Distance Measurement

We examine the results of filtering the two-dimensional Gaussian distribution

$$\mathrm{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1.0 & 0.5\\0.5 & 1.0\end{bmatrix}\right) \tag{19}$$

based on a distance measurement

$$\boldsymbol{y} = h(\underline{\boldsymbol{x}}) + \boldsymbol{v} = \sqrt{\boldsymbol{x}_1^2 + \boldsymbol{x}_2^2} + \boldsymbol{v} \tag{20}$$

with additive Gaussian noise $\boldsymbol{v}$. This problem results in posterior distributions reminiscent of two bananas. In this example, the measurement was set to $y = 2.0$ and the standard deviation of the measurement noise was set to 0.1. Initially, 100 samples were drawn from the prior Gaussian distribution and processed as in the EnGMF procedure. Fig. 4 shows the prior distribution, the estimated posterior GM and the reapproximation by samples for the two methods "LCD + Red" (top) and "PCD" (bottom). Similar to the example in Fig. 3 the samples drawn with "PCD" cover the posterior density more evenly than the sample drawn with "LCD + Red".

C. Nonlinear Dynamic System

We now apply the proposed filter to a difficult nonlinear filtering problem. We want to estimate the Lorenz96 system [13], which is often used as a reference because of its chaotic nature. To ensure that the system actually shows chaotic behavior, the forcing constant F has to be set accordingly. In

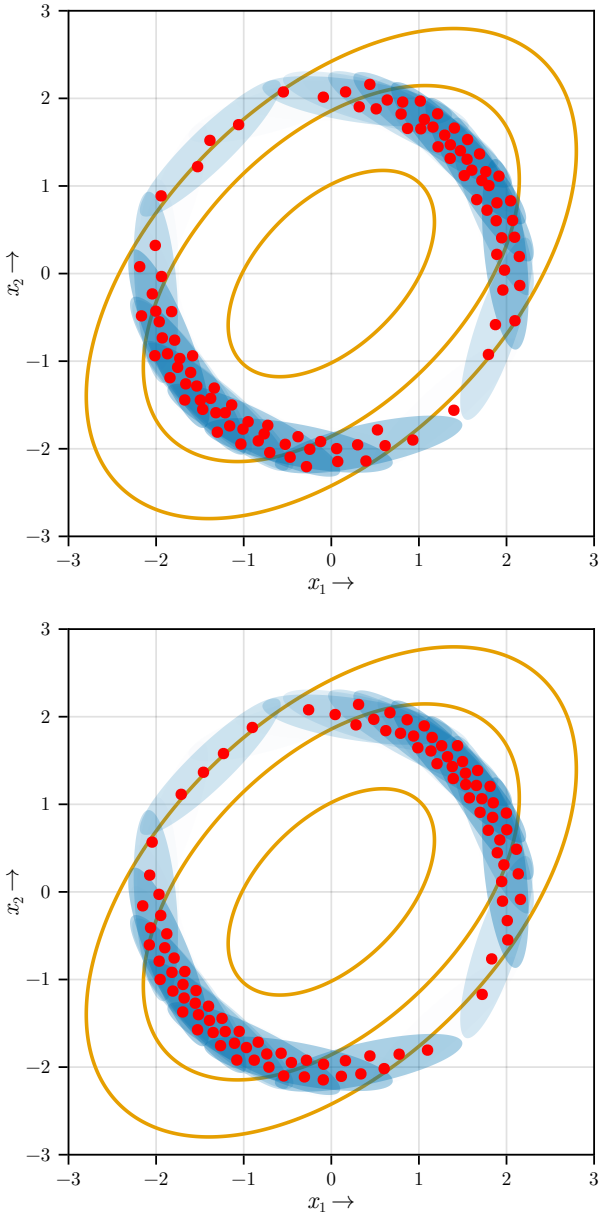

Fig. 4: Prior distribution (in yellow) and GM (in blue) and sample approximation (in red) of the posterior density for two approximation methods. (top) "LCD + Red" and (bottom) "PCD".

our examples it was therefore set to $F = 12$. We simulated 30 time steps of the four-dimensional and of the ten-dimensional Lorenz96 system with a time step of 0.01 and added Gaussian system noise with variance 0.01 in each dimension. For our experiments, 50 different initial positions were drawn from a standard normal distribution and the according trajectory was simulated.

The estimation itself was set up as a tracking problem based on range measurements. A measurement station was placed on each axis of the coordinate system at a distance of 5 units from the origin. Each station provides a measurement of its distance to the true target position corrupted by zero-mean additive Gaussian noise with variance 0.09 at each time step. The estimation was carried out for each of the 50 generated trajectories with four different filters and the Root Mean Square Error (RMSE) was calculated. We compared a particle filter with $10,000$ particles with three different variants of the EnGMF employing different methods for converting posterior GMs back to samples. These are the two methods "PCD" and "LCD+Red", that were already discussed in the previous examples and the solution based on optimal transport "OT" discussed in [18]. For all EnGMFs, 50 components were used and a Kernel density estimation with Silverman's rule-of-thumb was carried out before the filter step. For "PCD", 400 randomly sampled directions were used as projection direction. In the "LCD+Red" approach, 20 LCD samples were placed at each component and then reduced back to 50 samples. In the optimal transport based method, the calculation of the transport plan was iterated 10 times according to [7].

The results in Fig. 5 show that in four dimensions all the filters are able to track the target and give similar results. In ten dimensions, the particle filter shows worse results than the optimization-based EnGMFs. The optimal transport-based method initially performs well, but seems to lose the target after some time. Both optimization-based EnGMFs, i.e., "PCD" and "LCD+Red", successfully track the target and give very similar results.

VI. Conclusions

We proposed a new version of the EnGMF, where sampling from the posterior GMs is performed by systematically minimizing a distance measure. A new distance measure based on PCDs from [8] is used, which allows direct sampling from an entire GM without resorting to sampling from individual components first. This makes the sampling process simpler to implement and faster.

Usability of the EnGMF is strongly increased to its previous variants employing different sampling techniques. Compared to previous sampling techniques it shows higher quality per sample than random sampling, is faster than individual component sampling with subsequent reduction [6], and copes with higher dimensions compared to transport-based reduction [7].

VII. Outlook

In future work, we will conduct more complex simulations and include more methods for comparison. We will also systematically investigate the effect of different local filters such as the Unscented Kalman Filter (UKF) instead of the EKF on system performance. In addition, we will include different schemes for calculating the posterior weights for the GMs resulting from the locals filters.

References

[1] Jeffrey L. Anderson and Stephen L. Anderson. "A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts". *Monthly Weather Review* 127.12 (Dec. 1, 1999), pp. 2741–2758.

[2] Dalton Durant, Uwe D. Hanebeck, and Renato Zanetti. "Good Things Come to Those Who Weight! Improving Gaussian Mixture Weights for Cislunar Debris Tracking". *Proceedings of the AAS/AIAA Space Flight Mechanics Meeting, Kaua'i, Hawaii*. 25. Kaua'i, Hawaii, Jan. 2025, p. 309.

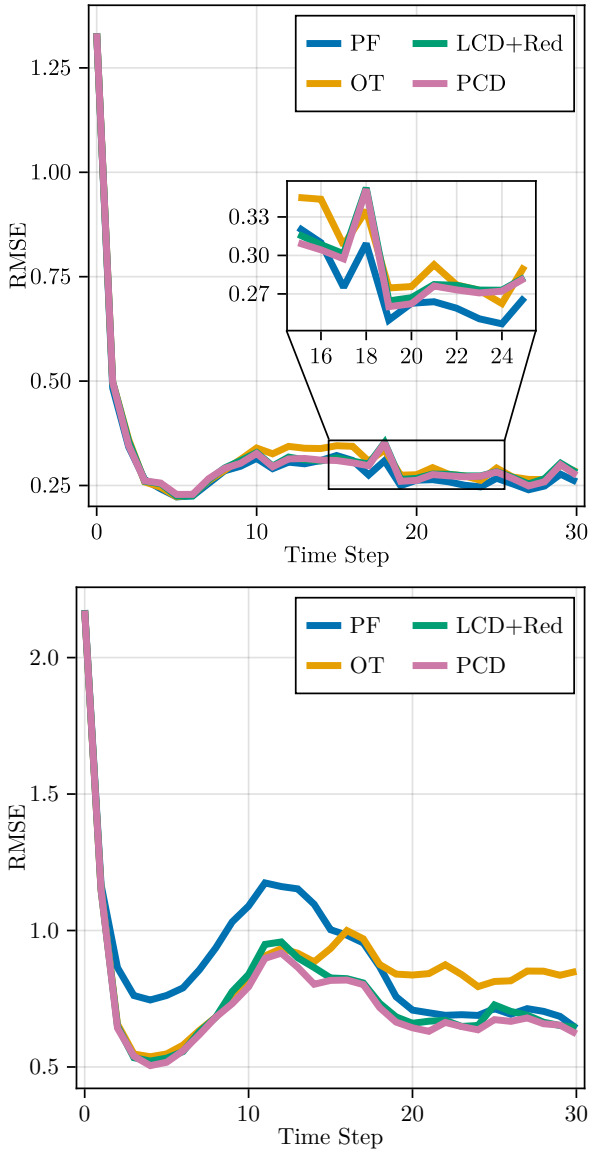

Fig. 5: Comparison of different resampling schemes for the EnGMF applied to a tracking problem based on the Lorenz96 system and range measurements. (top) four-dimensional system and (bottom) ten-dimensional system.

[3] Daniel Frisch. "Transformable Deterministic Sampling". Dissertation. Karlsruhe Institute of Technology (KIT), Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Referent: U. D. Hanebeck, Korreferent: S. Maskell, 2025. 301 pp.

[4] Daniel Frisch and Uwe D. Hanebeck. "Deterministic Gaussian Sampling With Generalized Fibonacci Grids". *Proceedings of the 24th International Conference on Information Fusion (Fusion 2021)*. Sun City, South Africa, Nov. 2021.

[5] Igor Gilitschenski et al. "Deterministic Dirac Mixture Approximation of Gaussian Mixtures". *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*. Salamanca, Spain, July 2014.

[6] Felipe Giraldo-Grueso et al. "Optimal Grid Point Sampling for Point Mass Filtering". *Proceedings of the AAS/AIAA Space Flight Mechanics Meeting, Kaua'i, Hawaii*. Kaua'i, Hawaii, Jan. 2025.

[7] Felipe Giraldo-Grueso et al. "Optimal Transport as a Reduction Technique for Deterministic Nonlinear Filtering". *Proceedings of the 28th International Conference on Information Fusion (FUSION 2025)*. Rio de Janeiro, Brazil, July 2025, pp. 1–8.

[8] Uwe D. Hanebeck. "Deterministic Sampling of Multivariate Densities based on Projected Cumulative Distributions". *Proceedings of the 54th Annual Conference on Information Sciences and Systems (CISS 2020)*. Princeton, New Jersey, USA, Mar. 2020.

[9] Uwe D. Hanebeck. "Optimal Reduction of Multivariate Dirac Mixture Densities". *at – Automatisierungstechnik* 63.4 (Apr. 2015), pp. 265–278.

[10] Uwe D. Hanebeck and Olga Feiermann. "Progressive Bayesian Estimation for Nonlinear Discrete-Time Systems: The Filter Step for Scalar Measurements and Multidimensional States". *Proceedings of the 2003 IEEE Conference on Decision and Control (CDC 2003)*. Maui, Hawaii, USA, Dec. 2003, pp. 5366–5371.

[11] Uwe D. Hanebeck, Marco F. Huber, and Vesa Klumpp. "Dirac Mixture Approximation of Multivariate Gaussian Densities". *Proceedings of the 2009 IEEE Conference on Decision and Control (CDC 2009)*. Shanghai, China, Dec. 2009.

[12] Marco Huber, Dietrich Brunn, and Uwe D. Hanebeck. "Closed-Form Prediction of Nonlinear Dynamic Systems by Means of Gaussian Mixture Approximation of the Transition Density". *Proceedings of the 2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006)*. Heidelberg, Germany, Sept. 2006, pp. 98–103.

[13] Dirk Leendert van Kekem. "Dynamics of the Lorenz-96 Model: Bifurcations, Symmetries and Waves". PhD thesis. Groningen: Rijksuniversiteit Groningen, 2018.

[14] Gerhard Kurz, Igor Gilitschenski, and Uwe D. Hanebeck. "Unscented von Mises–Fisher Filtering". *IEEE Signal Processing Letters* 23.4 (Apr. 2016), pp. 463–467.

[15] Kailai Li, Florian Pfaff, and Uwe D. Hanebeck. "Hyperspherical Deterministic Sampling Based on Riemannian Geometry for Improved Nonlinear Bingham Filtering". *Proceedings of the 22nd International Conference on Information Fusion (Fusion 2019)*. Ottawa, Canada, July 2019.

[16] Tiancheng Li, Tariq Pervez Sattar, and Shudong Sun. "Deterministic Resampling: Unbiased Sampling to Avoid Sample Impoverishment in Particle Filters". *Signal Processing* 92.7 (July 1, 2012), pp. 1637–1645.

[17] Andrey A Popov and Renato Zanetti. "Ensemble Gaussian Mixture Filtering with Particle-localized Covariances". *2023 26th International Conference on Information Fusion (FUSION)*. 2023 26th International Conference on Information Fusion (FUSION). June 2023, pp. 1–7.

[18] Andrey A. Popov and Renato Zanetti. *Deterministic Optimal Transport-based Gaussian Mixture Particle Filtering for Verifiable Applications*. Jan. 28, 2025. URL: http://arxiv.org/abs/2501.17302 (visited on 06/16/2025). Pre-published.

[19] Patrick Ruoff, Peter Krauthausen, and Uwe D. Hanebeck. "Progressive Correction for Deterministic Dirac Mixture Approximations". *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*. Chicago, Illinois, USA, July 2011.

[20] Bernard W. Silverman. *Density Estimation for Statistics and Data Analysis*. Boca Raton: Routledge, 2018. 176 pp.

[21] Peter Jan van Leeuwen et al. "Particle Filters for High-Dimensional Geoscience Applications: A Review". *Quarterly Journal of the Royal Meteorological Society* 145.723 (2019), pp. 2335–2365.

[22] Sehyun Yun, Renato Zanetti, and Brandon A. Jones. "Kernel-Based Ensemble Gaussian Mixture Filtering for Orbit Determination with Sparse Data". *Advances in Space Research* 69.12 (June 15, 2022), pp. 4179–4197.