Nonlinear Fusion of Multi-Dimensional Densities in Joint State Space

Vesa Klumpp and **Uwe D. Hanebeck**
Intelligent Sensor-Actuator-Systems Laboratory (ISAS),
Institute for Anthropomatics,
Universität Karlsruhe (TH), Germany.
klumpp@ira.uka.de, uwe.hanebeck@ieee.org

Abstract – *Nonlinear fusion of multi-dimensional densities is an important application in Bayesian state estimation. In the approach proposed here, a joint density over all considered densities is build, which is then approximated by means of a Dirac mixture density by partitioning the joint state space into regions that are represented by single Dirac components. This approximation procedure depends on the nonlinear fusion model and only areas relevant to this model are considered. The processing in joint state space has advantages, especially when fusing Dirac mixture densities. Within this approach, degeneration can be avoided and even densities without mutual support can be combined. Thus, this approach gives an alternative to multiplication of Dirac mixtures with a likelihood, as used in the particle filter. Furthermore, a nonlinear Bayesian estimator with filter and prediction step can be formulated, which is able to cope with both discrete and continuous densities.*

Keywords: Bayesian estimation, Dirac mixtures, Density approximation

1 Introduction

In Bayesian state estimation, the internal system state $\underline{x} \in \mathbb{R}^N$ is identified using measurements, which are disturbed by noise. Often, the relationship between the measurements and the system state is nonlinear, for example in localization, speech processing, or data fusion in sensor networks.

The estimated state is given by a probability density function, representing the knowledge and certainty about the true system state. In many implementations of Bayesian estimators, continuous densities are considered. Especially Gaussian or Gaussian mixture densities are favored because of their useful features, like the closedness regarding linear transformations and the small number of parameters needed. Their disadvantage is that for nonlinear transformations, these densities have to be re-approximated in order to limit representational complexity or to allow processing at all.

In this paper, we consider the case of processing Dirac mixture densities. Their main advantage is the fact that Dirac mixture densities are simple to propagate through complex nonlinear models, which makes them practical for nonlinear state estimation. These densities are represented by a weighted sum of Dirac components, or samples. They are usually simple to process, although many of them are needed. Another property, which is discussed later, is their selective support, which can be harmful when processing them by means of a Bayesian filter step, where densities are multiplied. Thus, a different approach for fusion of Densities has to be pursued, which is based on an approximation of a joint density over all considered densities.

Novelties

The novelties of this paper are the generalizations of our previous work [1] for Bayesian fusion with Dirac mixture densities in joint state space. Now, several multi-dimensional random variables and nonlinear fusion constraints are considered. The problem of the approximation of the integral, which stems from the fusion constraint, is efficiently solved by an approximation. Furthermore, this framework allows the application as a Bayesian estimator for both filter and prediction step by the same principle.

2 Related Work

Several ways for approximating continuous probability density functions by means of Dirac mixture densities exist. The most simple approach is random sampling [2], where a number of samples is drawn independently from a given density function. A better representation quality is obtained by using deterministic Dirac mixture approximation algorithms. Here, distance measures between the true density and the approximation are employed [3, 4]. These algorithms can be classified by two criteria: batch/sequential and optimal/suboptimal algorithms. Batch approaches return the approximating Dirac mixture density as a final result, wheras sequential algorithms iteratively refine the approximation by adding components sequentially [5]. Optimal algorithms return the best possible result under the given distance measure [6], whereas suboptimal algorithms apply approximations or heuristics in order to reduce computations costs [5].

In the framework of Bayesian state estimation, often Dirac mixture densities, or sample sets, are used. The most popular approach is the particle filter [7], for which a vast number of extensions exists. Other estimation algorithms with different kinds of density representations are available

that make use of a sampled representation for nonlinear processing. Here, samples are drawn from a continuous density function, i.e., a Gaussian density, which are propagated through a nonlinear system model. After this, a continuous density representation is derived from the resulting sample set. Algorithms of this type are the Unscented Kalman filter [8] or the Gauss filter [9].

A different problem occurs when combining several Dirac mixture densities by means of the Bayesian filter step. Here, the discrete type of the densities can be problematic, due to different support. A solution to overcome this is approximating one density by a continuous form, e.g., by convolution of one Dirac mixture density with a Gaussian kernel. This leads to a continuous density function, which can be combined with the second Dirac mixture density by simple pointwise multiplication [10]. A problem herein is that the kernel size has to be known in advance. A completely different approach is pursued in [1]. Here, a joint density over all involved random variables is considered. This density is reapproximated by means of a Dirac mixture, in such a way that the resulting density of the combination can be obtained by a projection. This special approach is considered here and explained in the following sections.

3 Problem Formulation

A Dirac mixture density is defined as

$$f(\underline{x}) = \sum_{i=1}^{L} w^i \cdot \delta\left(\underline{x} - \underline{\xi}^i\right) \ ,$$

with the weights

$$w^i \geq 0 \ \forall i \in \{1, \ldots, L\} \text{ and } \sum_{i=1}^{L} w^i = 1$$

and Dirac positions $\underline{\xi}^i$. The N-dimensional Dirac delta function is defined by the product

$$\delta(\underline{x}) = \prod_{i=1}^{N} \delta\left(x^{(i)}\right) \ , \ \underline{x} = [x^{(1)}, \ldots, x^{(N)}]^T$$

of one-dimensional Dirac distributions. Note, that throughout this paper, vectors $\underline{x}$ are underlined and random variables $\boldsymbol{x}$ are bold face.

We consider the Bayesian fusion of two or multiple random vectors, $\underline{\boldsymbol{x}}_1, \ldots, \underline{\boldsymbol{x}}_n$. The fusion of density functions is usually performed according to the Bayes equation. Here, the normalized product of two densities has to be calculated. If both densities are given as Dirac mixtures, several problems arise. One problem is that both densities usually have no common support. In this case, the resulting density is zero, as visualized in Figure 1, which is useless for state estimation. Another problem is that simple pointwise multiplication of two Dirac mixture densities does not correspond to the multiplication of two continuous densities. This is shown by the following example.

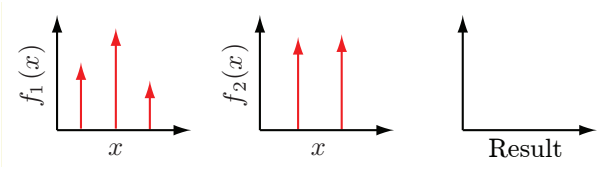

Figure 1: Pointwise multiplication of two Dirac mixture densities $f_1(x)$ and $f_2(x)$. The result is zero because they have no mutual support.

Example 1 (Multiplication of two Dirac mixtures)
We consider two identical Dirac mixture densities f_1, f_2 with uniform weights, which approximate a Gaussian density $\widetilde{f}$ with variance $\sigma^2 = 1$. The product of both Dirac mixture densities $f_1 \cdot f_2$ has again equal weights and the Dirac positions did not change. Thus, the product is equal to the prior densities $f_1 \cdot f_2 = f_1 = f_2$, which conflicts with the multiplication of Gaussians. The product of two Gaussian densities with variance 1 is a Gaussian density with variance $1/2$.

These problems lead to a different interpretation and processing of Dirac mixture densities. Instead of regarding a Dirac mixture density as the corresponding representation of a random variable, a slightly different approach is taken. Here, we assume that a Dirac mixture density f is an approximation of an underlying true, continuous density $\widetilde{f}$ [1]. Hence, we need to find a processing step, which generates the approximation of the true posterior density function by means of a Dirac mixture, given the approximations of the prior densities only. For Example 1, the result would be a Dirac mixture approximation of the posterior product $\widetilde{f} \cdot \widetilde{f}$, without knowledge about the true densities.

4 Processing in Joint State Space

In this section, the idea of processing multiple densities by means of Bayesian fusion in joint state space is explained. The general model equation for the fusion of random vectors can be expressed in terms of

$$g(\underline{\boldsymbol{x}}_1, \ldots, \underline{\boldsymbol{x}}_n) = 0 \ , \tag{1}$$

with n random vectors and an arbitrary function $g(\cdot)$, which defines the combination of random vectors.

In this approach, the joint state space over all considered random variables is spanned. The density over the joint state space $\widetilde{f}_J(\underline{z}) = \widetilde{f}_J(\underline{x}_1, \ldots, \underline{x}_n)$ is given in case of independent random variables as

$$\widetilde{f}_J(\underline{x}_1, \ldots, \underline{x}_n) = \prod_{i=1}^{n} \widetilde{f}(\underline{x}_i) \ ,$$

where $\widetilde{f}(\underline{x}_i)$ are the densities of the N_i-dimensional random variables $\underline{\boldsymbol{x}}_i$. N_J is the dimensionality of the joint density $\widetilde{f}_J$.

In order to perform the fusion step, the joint density has to be evaluated on points where Equation (1) holds. Let i denote the posterior subspace of the fusion result. The posterior density can be derived by integrating over all subspaces

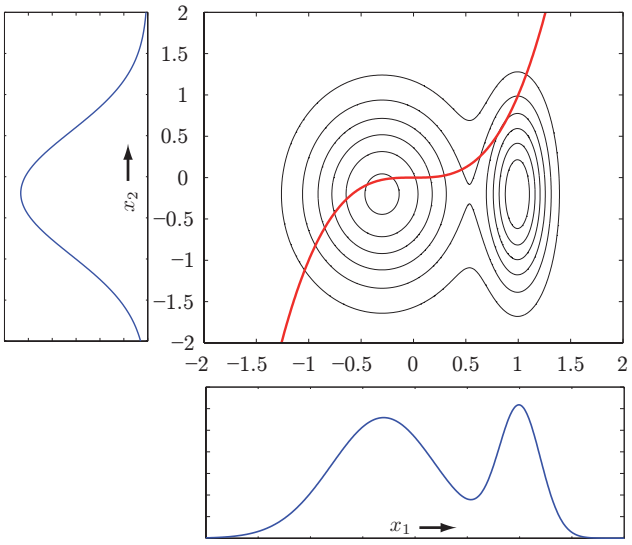

Figure 2: Fusion of two densities in joint state space. The resulting density is on the constraint $x_2 = x_1^3$ (red).

except i according to

$$\widetilde{f}_p(\underline{x}_i) = \int_{g(\underline{x}_1,\ldots,\underline{x}_n)=0} \widetilde{f}_J(\underline{z})\,\mathrm{d}\underline{x}_n \cdots \mathrm{d}\underline{x}_{i+1}\,\mathrm{d}\underline{x}_{i-1} \cdots \mathrm{d}\underline{x}_1\,, \quad (2)$$

where the joint density is evaluated at points, where equation (1) holds. All in all, the model equation is regarded as a constraint in the joint state space.

Example 2 (Nonlinear Fusion Constraint)
The joint state space of two Gaussian mixture densities $f(x_1)$ and $f(x_2)$ is shown in Figure 2. The fusion constraint $x_2 = x_1^3$ is marked red.

With the representation (2) of the posterior density, the problem of incompatible supports still exists, because it is not guaranteed that the joint density $\widetilde{f}_J$ has support on the constraint when dealing with Dirac mixtures. In order to overcome this, an approximation of the joint density f_J with support on the constraint has to be found. In this paper, $f_J(\underline{z})$ is a Dirac mixture density, which can be transformed into a posterior Dirac mixture density by simple projection.

5 Approximation in Joint State Space

5.1 General Algorithm

The Dirac mixture approximation of the true joint density $\widetilde{f}_J$ presented here is a modification of [5]. Every Dirac component is an approximation of the true density over a bounded region. In order to refine the approximation, the regions are split and the new regions are reapproximated by new Dirac components. This allows a sequential and local refinement of the approximation. Dirac components that have no influence on the posterior density, i.e., regions that do not interfere with the constraint, can be eliminated. When the desired number of Dirac components is reached, the posterior density f_p is calculated. A detailed description of the algorithm is given in the following.

Approximation of Region
A rectangular axis-aligned region is defined by all points $\underline{z}$, for which

$$\underline{l} \leq \underline{z} \leq \underline{u}, \quad \underline{l}, \underline{u} \in \mathbb{R}^{N_J}$$

holds. Here, the relation $\leq$ holds, if the relation $\leq$ on scalar values holds for all N_J dimensions. This region is denoted by $[\underline{l}, \underline{u}]$. One Dirac component with position $\underline{\xi}_i$ and weight w^i represents the approximation of the true density $\widetilde{f}_J(\underline{z})$ in this region. The weight

$$w^i = \int_{l^{(1)}}^{u^{(1)}} \cdots \int_{l^{(N_J)}}^{u^{(N_J)}} \widetilde{f}_J(\underline{z})\,\mathrm{d}z^{(N_J)} \cdots \mathrm{d}z^{(1)}$$

is given by the probability over the region.

The Dirac position is determined by considering all one-dimensional marginals over the region. These so-called submarginals are defined as

$$\widetilde{f}_J^{\,i}(x^{(i)}) = \int_{l^{(1)}}^{u^{(1)}} \cdots \int_{l^{(i-1)}}^{u^{(i-1)}} \int_{l^{(i+1)}}^{u^{(i+1)}} \cdots \int_{l^{(N_J)}}^{u^{(N_J)}} \widetilde{f}_J(\underline{z})\;\mathrm{d}z^{(N_J)} \cdots \mathrm{d}z^{(i+1)}\,\mathrm{d}z^{(i-1)} \cdots \mathrm{d}z^{(1)}$$

for $l^{(i)} \leq z^{(i)} \leq u^{(i)}$ and 0 elsewhere, for dimension i and region $[\underline{l}, \underline{u}]$. For continuous submarginals, the optimal position is the median $m^{(i)}$, which minimizes the Cramér-von Mises distance [11]

$$D^i = \int_{l^{(i)}}^{u^{(i)}} \left(\widetilde{F_J}^{\,i}(x^{(i)}) - F_J^i(x^{(i)})\right)^2 \mathrm{d}x^{(i)}$$

of the cumulative distributions over the submarginals. For Gaussian mixture densities $\widetilde{f_J}^{\,i}$, the median cannot be found analytically. Thus, a simple numerical solver for the problem of median finding

$$\widetilde{F_J}^{\,i}(m^{(i)}) \stackrel{!}{=} \frac{1}{2}\left(\widetilde{F_J}^{\,i}(l^{(i)}) + \widetilde{F_J}^{\,i}(u^{(i)})\right)\,,$$

like the procedure of nested intervals [12], can be employed.

If the submarginal is a Dirac mixture, there is no distinct point that exactly halves the probability mass. Therefore, a linear interpolation for the cumulative submarginal $\widetilde{F_J}^{(i)}$ can be performed. The interpolation approximates the submarginal by piecewise uniform parts.

Remark 1 (Interpolation of Dirac Mixture)
Let $\widetilde{f}(x)$ be a Dirac mixture with L components, weights w^i, and Dirac positions ξ^i. We assume that the Dirac positions are ordered, i.e., $\xi^1 < \ldots < \xi^L$, and the first and last Dirac components are given by $\xi^1 = l^{(i)}, w^1 = 0$ and

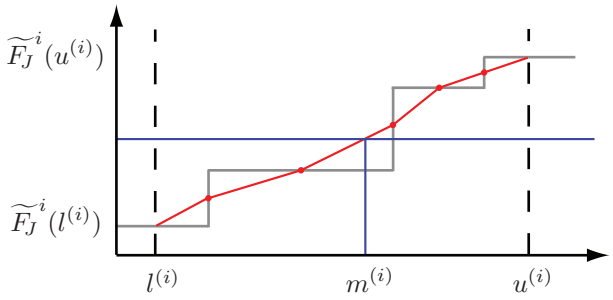

Figure 3: Interpolation of Heaviside step functions by straight lines for median computation.

$\xi^L = u^{(i)}, w^L = 0$. *Then, the interpolation of the Dirac mixture density is given by*

$$\overline{f}(x) = \begin{cases} \dfrac{w^i}{\xi^{i+1} - \xi^i}\,, & \xi^i \le x < \dfrac{\xi^i + \xi^{i+1}}{2} \\ \dfrac{w^{i+1}}{\xi^{i+1} - \xi^i}\,, & \dfrac{\xi^i + \xi^{i+1}}{2} \le x < \xi^{i+1} \end{cases} .$$

With this continuous representation of the submarginal, the median can be determined. The interpolation is depicted in Figure 3.

Splitting

The splitting step is performed in order to improve the approximation quality at regions of high interest. Improvement is achieved by inserting additional Dirac components. For this, the Dirac component of the selected region is discarded and the region is split into two new regions. The splitting is performed along axis i

$$i = \arg\max \left\{ u^{(i)} - l^{(i)}, i = 1, \ldots, N_J \right\}$$

with maximum extent. Note, that this approach can be modified to cope with different scales, e.g., position and angle components in the state vector, by rescaling the problem into the unit hyper-cube $[0,1]^{N_J}$.

The splitting intersects the old Dirac component, and thus, the new regions $[\underline{l}_1, \underline{u}_1]$, $[\underline{l}_2, \underline{u}_2]$ are given by

$$\begin{aligned} \underline{l}_1 &= \underline{l}\,, \\ \underline{u}_1 &= \left[u^{(1)}, \ldots, u^{(i-1)}, \xi^{(i)}, u^{(i+1)}, \ldots, u^{(N_J)}\right]^T , \\ \underline{l}_2 &= \left[l^{(1)}, \ldots, l^{(i-1)}, \xi^{(i)}, l^{(i+1)}, \ldots, l^{(N_J)}\right]^T , \\ \underline{u}_2 &= \underline{u}\,. \end{aligned}$$

These regions are then approximated by Dirac components.

With certain submarginals, it is not possible to give a usable splitting. This is the case, when the particular submarginal $\widetilde{f}_J^i$ consists of one Dirac component only. If this occurs, another dimension has to be chosen or a different region should be split. After splitting, it is recommended to check the regions on common points with the constraint. If they do not comply with the fusion constraint, they can be discarded immediately. This problem is discussed further in Section 5.2.

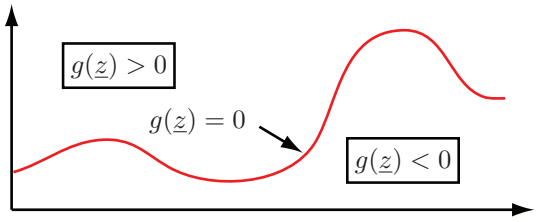

Figure 4: The constraint g divides the joint state space into a positive and a negative part. On the constraint, g is zero.

Region Selection

The splitting candidate for the next iteration is determined by the region selection step. Different approaches for determining the next region are possible. In [5], regions are arranged in a tree, in which the regions are selected in a disperse and symmetrical way, in order to give a regular approximation of the complete state space. Other approaches include the selection of the region with highest enclosing probability, maximum extent, or maximum volume. The whole iteration cycle is continued, until the desired approximation quality is reached, the number of desired Dirac components is generated, or no further splittings are possible. Due to simplicity, the approach taken here choses regions with maximum weight.

5.2 Constraint Evaluation

Only relevant regions, which contain parts of the constraint, have to be considered for the fusion result. The problem arising in the given approximation algorithm is that the constraint has to be evaluated for a set, i.e., every point in the region $[\underline{l}, \underline{u}]$. Thus, depending on the fusion constraint g, we need to decide, whether the set

$$\mathcal{A} = \left\{ \underline{z} : g(\underline{z}) = 0,\ \underline{l} \le \underline{z} \le \underline{u},\ \underline{z} \in \mathbb{R}^{N_J} \right\}$$

is empty or not, for the given region. In the special case of continuous, piecewise monotone functions g, the regions are simple to check, if it is possible to decide on which side of the constraint a given point $\underline{z}$ is. This property is shown in Figure 4. Here, g returns positive or negative values, depending on whether a point is located above or below the constraint. With this representation of the constraint, only certain points have to be considered.

Remark 2 (Points for Constraint Evaluation)

The equation $g(\underline{z}) = 0$ defines a manifold in joint state space $\mathbb{R}^{N_J}$. The point set $\mathcal{V}$ consists of the vertices of the hyper-cuboid, i.e., the corners of the region, and points that are relevant to local extremal values of the manifold along the axes. For every local extremal point $\underline{x}^i$ of the manifold,

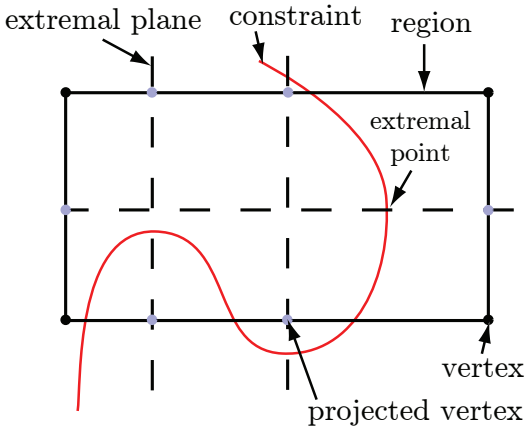

Figure 5: Evaluation of constraint. If the vertices and projected vertices lie on both sides of the constraint, then the constraint is met. The vertices are projected onto the extremal planes, which are defined by the extremal points of the constraint.

a hyper-plane p^i perpendicular to the corresponding axis is generated. The vertices $\mathcal{V}$ and the projections of $\mathcal{V}$ onto the hyper-planes through the extremal points define the set $\mathcal{X}$, i.e.,

$$\mathcal{X} = \mathcal{A} \cup \left(\bigcup_i \{\mathcal{A} \downarrow p^i\} \right) .$$

The local extremal points of the constraint, the vertices of the region, and their projections are depicted in Figure 5. With the point set $\mathcal{X}$, intersection can be checked for the corresponding region.

Remark 3 (Constraint Evaluation for Region)
If $g(\underline{x})$ has different signs for all points $\underline{x} \in \mathcal{X}$, i.e.,

$$\exists\, \underline{x}_1, \underline{x}_2 \in \mathcal{X} : g(\underline{x}_1) < 0 < g(\underline{x}_2) ,$$

then there exists an intersection between the constraint and the region $[\underline{l}, \underline{u}]$.

5.3 Result: Posterior Density

After Dirac approximation of the joint density $\widetilde{f}_J$, the final fusion result is calculated. After approximation, the resulting Dirac components lay in the neighbourhood of the constraint and have to be mapped onto the lower dimensional subspace of the posterior density f^p. The resulting Dirac mixture, i.e., the positions and weights of all "active" components in joint state space, is given by

$$f_J(\underline{z}) = \sum_{i=1}^{L^p} w^i \delta(\underline{z} - \underline{\xi}^i) . \tag{3}$$

This density has to be mapped to a posterior density

$$f^p(\underline{x}) = \sum_{i=1}^{L^p} w^{p,i} \delta(\underline{x} - \underline{\xi}^{p,i}) ,$$

with weights $w^{p,i}$, which sum up to 1, and Dirac positions $\underline{\xi}^{p,i}$.

Posterior Weights
The components, especially the weights w^i, in (3) cannot be used directly, because the approximation considers complete regions in joint state space, whereas the fusion result lies on the constraint only. This additional extent and probability mass has to be considered.

Another problem is that an approximating Dirac component $\underline{\xi}^i$ in joint state space usually does not lie on the constraint and thus, the posterior density would still be zero. In order to overcome this, the probability over the region $[\underline{l}^i, \underline{u}^i]$ is assumed to be uniformly distributed, i.e.,

$$f_J(\underline{z}) \approx \frac{w^i}{\prod_{n=1}^{N_J} u^{(i)} - l^{(i)}} \quad , \underline{l}^i \leq \underline{z} \leq \underline{u}^i . \tag{4}$$

This can be assumed if the regions are small and the true underlying densities are smooth to a certain degree.

The resulting posterior density is then obtained by evaluating (4) over the constraint $\underline{g}$. In the case of posterior Dirac mixture densities, the weight $w^{p,i}$ is calculated according to the integral

$$w^{p,i} = c \cdot \int_{\mathcal{A}^i} \frac{w^i}{\prod_{n=1}^{N_J} u^{(i)} - l^{(i)}} \, \mathrm{d}\underline{z} , \tag{5}$$

with the set

$$\mathcal{A}^i = \{\underline{z} : g(\underline{z}) = 0\} \cap [\underline{l}^i, \underline{u}^i]$$

and a normalizing factor $c = 1/\sum_{i=1}^{L^p} w^{p,i}$. The integral (5) is usually difficult to solve, especially the calculation of the intersection $\mathcal{A}^i$ is cumbersome for high-dimensional joint state spaces and $\mathcal{A}^i$ does not have to be a connected set. For the special case of two scalar densities $f(x)$, $f(y)$ and the fusion constraint $y = x$, an analytic expression for the posterior weight can be specified [1].

For multi-dimensional densities and nonlinear fusion constraints, approximation methods for the posterior weight $w^{p,i}$ have to be found. The approximation is based on substituting the integral (5) by a formula, which can be easily and efficiently evaluated and maintains the convergence of the posterior Dirac mixture f^p against the true posterior density $\widetilde{f}^P$ with growing number L^p of components. The approximate weight is given by

$$w^{p,i} \approx c \cdot w^i \cdot \frac{\sqrt{\sum_{i=1}^{N_J} \left(u^{(i)} - l^{(i)}\right)^2}}{\prod_{i=1}^{N_J} \left(u^{(i)} - l^{(i)}\right)} ,$$

with a normalization constant $c = \sum_{i=1}^{L^p} w^{p,i}$. For a growing number of iterations, the region $[\underline{l}^i, \underline{u}^i]$ shrinks and the weight $w^{p,i}$ converges towards the true weight. Note, that for a large numer of Dirac components $L^p \to \infty$, the posterior weight $w^{p,i}$ is equal to w^i.

Posterior Dirac Positions

For the posterior Dirac position, a mapping of the position $\underline{\xi}^i$ in joint state space onto $\underline{\xi}^{p,i}$ on the posterior subspace has to be found. Here, also different methods with varying approximation quality exist.

The optimal solution for the posterior Dirac position according to this framework is the optimal approximation of the density on the constraint with one Dirac component on the constraint within a region. The approximated density in the joint state space is given as piecewise uniform densities. If the approximation (4) is evaluated on the constraint only, the density

$$f(\underline{z}) \approx \begin{cases} \dfrac{w^i}{\prod_{n=1}^{N_J} u^{(i)} - l^{(i)}} & , \ \underline{z} \in \mathcal{A}^i \\ 0 & , \text{ elsewhere} \end{cases}$$

has to be approximated with one Dirac component that will be projected onto the posterior subspace $\underline{x}^p$. As mentioned above, the set $\mathcal{A}^i$ can be non-connected and difficult to evaluate for nonlinear fusion constraints. Thus, other approximation methods with less computational effort are employed.

The simplest way is an orthogonal projection onto $\underline{x}^p \in \mathbb{R}^N$, denoted as

$$\underline{\xi}^{p,i} = \underline{\xi}^i \downarrow \mathbb{R}^N \ ,$$

which is sufficient for regions of small extent and a large number of Dirac components L^p.

A tradeoff between both approaches, which will be used in the examples, is given in the following. Here, the Dirac component is first mapped onto the constraint and then projected onto the posterior subspace. The posterior Dirac position is then given by

$$\underline{\xi}^{p,i} = \left(\underline{\xi}^i \downarrow g\right) \downarrow \mathbb{R}^N \ .$$

This preserves the condition that the Dirac component must lie on the constraint before projection onto $\underline{x}^p$.

Example 3 (Projection onto constraint)

An example projection is shown in Figure 6. Here, a two-dimensional joint space is approximated on the constraint

$$0.1(x_1)^3 - 0.5(x_1)^2 + 1 - x_2 = 0$$

and the resulting Dirac components (blue) are mapped onto the constraint (red).

6 Fusion in Joint State Space

Now, with the introduced framework for processing density functions with a fusion constraint, a Bayesian estimator can be constructed. This estimator allows the processing of continuous (e.g., Gaussian mixture) and discrete (e.g., Dirac mixture) densities by the same principle. All occuring densities can be of both types, whereas the resulting density of the processing steps is always a Dirac mixture.

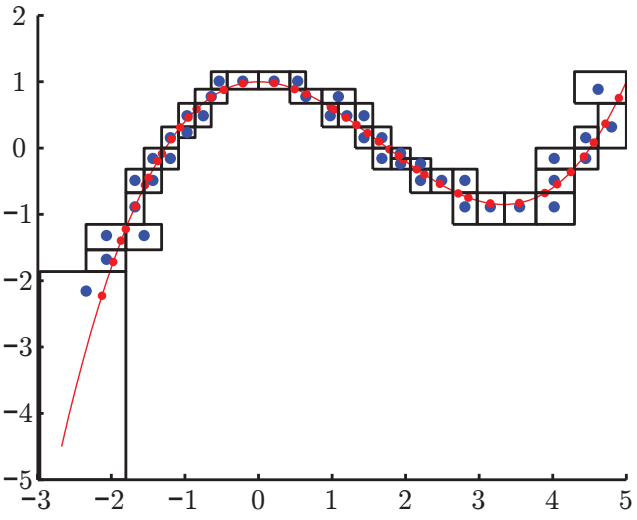

Figure 6: Approximation of the joint state space on the constraint.

Filter Step

For Bayesian state estimation, a measurement model

$$\hat{\underline{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{v}}_k) \tag{6}$$

is given, which is a special case of the general model (1). The mapping $\underline{h}_k$ describes the relationship between the state $\underline{x}_k \in \mathbb{R}^N$ and the measurement $\hat{\underline{y}}$ for discrete time steps k, which is disturbed by a noise term $\underline{\boldsymbol{v}}_k$. The estimated state is given by a density function $f^e(\underline{x}_k)$. The filter step is processed according to the Bayes law,

$$f^e(\underline{x}_k) = \frac{f(\underline{x}_k) \cdot f^L(\hat{\underline{y}}_k | \underline{x}_k)}{\int_{\mathbb{R}^N} f(\underline{\xi}) \cdot f^L(\hat{\underline{y}}_k | \underline{\xi}) \, \mathrm{d}\underline{\xi}} \ ,$$

the normalized product of prior density f^p and likelihood f^L, which depends on the measurement noise distribution, the measurement equation, and the measurement.

In order to apply the processing in Section 5, a suitable constraint function g has to be found.

Definition 1 (Constraint for Bayesian Filter Step)

The constraint function g for a measurement model (6) is given by

$$g(\underline{x}_k, \underline{v}_k) = \sum_{\substack{i=1 \\ i \neq j}}^{N} |h_k^{(i)}(\underline{x}_k, \underline{v}_k) - \hat{y}_k^{(i)}| + h_k^{(j)}(\underline{x}_k, \underline{v}_k) - \hat{y}_k^{(j)} \ ,$$

where the upper index (i) denotes the i-th element of the vectors.

The dimension j is used to decide whether a given point is above or below the constraint, as depicted in Figure 4.

The result of Bayesian fusion is the density $f^e(\underline{x}_k)$, and thus, the posterior density has to be a projection onto the

$\underline{x}_k$ subspace. As shown later, this approach prevents degeneration, even when the prior densities are given as Dirac mixtures and have no mutual support.

Prediction Step
The prediction step is performed according to the discrete time model

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{w}}_k) \,, \tag{7}$$

with the state estimate $\underline{\boldsymbol{x}}_k$ at time step k, a system noise term $\underline{\boldsymbol{w}}_k$, and a mapping $\underline{a}_k$ that updates the estimate for the next time step $k+1$. Note, for simplicity, any input $\underline{u}_k$ is omitted here. The densities are calculated according to the Chapman-Kolmogorov equation

$$f(\underline{x}_{k+1}) = \int_{\mathbb{R}^N} f^T(\underline{x}_{k+1}|\underline{x}_k) \cdot f(\underline{x}) \, \mathrm{d}\underline{x}_k \,,$$

where the transition density f^T depends on the model equation $\underline{a}_k$ and the noise term $\underline{\boldsymbol{w}}_k$.

In the joint state space framework, the prediction step is processed in the same way as the filter step. The joint density is approximated on a constrain, which is given analogously in the filter step.

Definition 2 (Constraint for Prediciton Step)
The constraint function g for the system model (7) is given by

$$g(\underline{x}_{k+1}, \underline{x}_k, \underline{w}_k) = \sum_{\substack{i=1 \\ i \neq j}}^{N} |a_k^{(i)}(\underline{x}_k, \underline{w}_k) - x_{k+1}^{(i)}| + a_k^{(j)}(\underline{x}_k, \underline{w}_k) - x_{k+1}^{(j)} \,.$$

Here, the joint state space is $3 \cdot N$-dimensional and spanned over the subspaces of $\underline{x}_k$, $\underline{w}_k$, and $\underline{x}_{k+1}$. Because no prior information about the distribution of $\underline{\boldsymbol{x}}_{k+1}$ should be regarded, $\underline{\boldsymbol{x}}_{k+1}$ is uniformly distributed. The resulting posterior Dirac mixture density is then projected onto the $\underline{x}_{k+1}$ subspace.

7 Examples

7.1 Degeneration Free Filtering

This example shows the use of the Bayesian filter step.The measurement equation is given as

$$\hat{y}_k = \boldsymbol{x}_k^3 + \boldsymbol{v}_k \,,$$

with a zero mean Gaussian distributed noise term $\boldsymbol{v}_k$ with variance 1. The true prior density is also Gaussian $\widetilde{f}(x_k) = \mathcal{N}(x_k - 0, 9)$. For this example, $\widetilde{f}$ is approximated by a Gaussian mixture with 8 equally weighted Dirac components. The measurement received is $\hat{y}_k = 4$. The prior Dirac mixture and the likelihood are visualized in the top of Figure 7. The estimated densities are visualized at the bottom of Figure 7. Here, the true density (blue, dashed) is compared to the particle filter (red) and the proposed approach in (red). It can be seen, that the particle filter suffers from degeneration. Here, only one component is used to represent the estimated distribution.

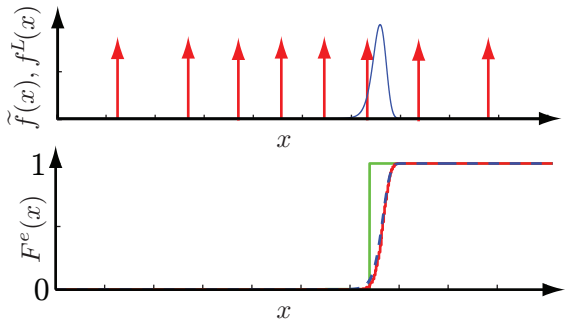

Figure 7: Illustration of degeneration. Top: Prior density (red) and likelihood (blue). Bottom: The true estimated density (blue, dashed), the reweighted particles of the particle filter (green) and the proposed approach (red) are shown.

7.2 Filter Step for Densities With Contradictory Information

In this example, Bayesian filtering of densities without mutual support is shown. The measurement equation is given by

$$\hat{y}_k = \boldsymbol{x}_k^2 + \boldsymbol{v}_k \,.$$

$\boldsymbol{v}_k$ is a noise term with the density function $f^v(x_k) = \mathcal{N}(x_k - 0, 0.2)$. The measurement processed is $\hat{y}_k = 5$. The prior density is an approximation of the Gaussian density $\widetilde{f}(x_k) = \mathcal{N}(x_k - 0, 0.2)$ with 8 equally weighted components. The prior density (red) and the likelihood (blue) are shown in the top of Figure 8.

At the bottom of Figure 8, fusion results are shown. The true distribution (blue, dashed) is compared to simple re-weighting of particles according to the likelihood (green) and to the proposed approach (red). Simple re-weighting does not change the support and is not able to represent the true result appropriately. Here, also degeneration occurs: only the outer left and right particles have relevant weights, whereas the remaining components have no impact to the result. The proposed approach is able to track the true fusion result with minor deviation. This error is due to the rough representation of the prior density by only 8 components.

8 Conclusion and Future Work

This paper presents a generalization of previous work. Now, a general framework for nonlinear processing of multi-dimensional density functions is considered. It is based on a density approximation in joint state space over all densities that shall be combined. Within the joint state space, a Dirac mixture approximation is performed, which ensures a result that is located on the nonlinear fusion constraint. This approximation partitions the joint state space into rectangular areas, which can be handled efficiently

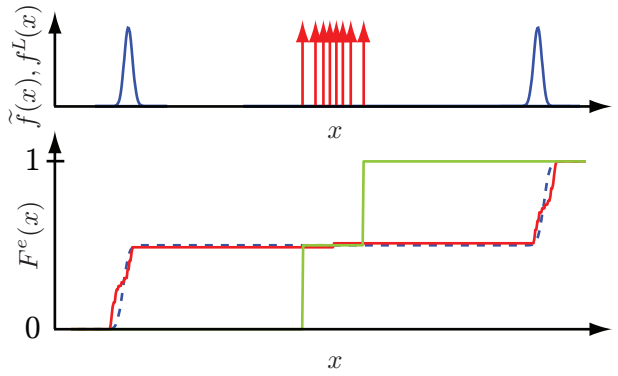

Figure 8: Illustration of the filtering. Top: Prior density (red) and likelihood (blue). Bottom: Fusion result after nonlinear measurement. The true fusion result (blue, dashed) is compared to simple re-weighting (green) and the proposed approach (red).

Depending on the application, this approximation algorithm allows a tradeoff between the approximation quality and the number of used components. The specialty of this processing is that even when the densities have no mutual support, a feasible result can be obtained. Additionally, the degeneration problem, where only few components of a Dirac mixture contribute to the density representation, can be avoided. This makes the contributed approach an alternative for the fusion of sampled densities with a continuous likelihood.

Future work includes the use of the given approach in the Sliced Gaussian Mixture Filter [13] in order to avoid degeneration. Another possible approach for the problem for the determination of the Dirac components on the constraint is interpolation of Dirac components. For a given approximation of the joint state space a Dirac mixture interpolation on the constraint could be found that minimizes certain criteria, e.g., maximum smoothness. This would allow the use of random sampling or offline calculated Dirac mixture approximations of continuous densities, like Gaussians, that would reduce the runtime of the algorithm.

References

[1] V. Klumpp and U. D. Hanebeck, "Direct Fusion of Dirac Mixture Densities using an Efficient Approximate Solution in Joint State Space," in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008.

[2] J. H. Kotecha and P. M. Djurić, "Gaussian Particle Filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2592–2601, 2003.

[3] U. D. Hanebeck and O. C. Schrempf, "Greedy Algorithms for Dirac Mixture Approximation of Arbitrary Probability Density Functions," in *Proceedings of the 2007 IEEE Conference on Decision and Control (CDC 2007)*, New Orleans, Louisiana, Dec. 2007, pp. 3065–3071.

[4] 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.

[5] V. Klumpp and U. D. Hanebeck, "Dirac Mixture Trees for Fast Suboptimal Multi-Dimensional Density Approximation," in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008.

[6] O. C. Schrempf and U. D. Hanebeck, "Recursive Prediction of Stochastic Nonlinear Systems Based on Dirac Mixture Approximations," in *Proceedings of the 2007 American Control Conference (ACC 2007)*, New York, New York, Jul. 2007, pp. 1768–1774.

[7] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[8] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[9] M. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, Seoul, Korea, Jul. 2008.

[10] L.-L. Ong, B. Upcroft, M. Ridley, T. Bailey, S. Sukkarieh, and H. Durrant-Whyte, "Consistent methods for Decentralized Data Fusion using Particle Filters," in *Proceedings of the 2006 IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006)*, Heidelberg, Germany, Sep. 2006.

[11] O. Ozturk and T. P. Hettmansperger, "Generalized Weighted Cramér–von Mises Distance Estimators," *Biometrika*, vol. 84, no. 2, pp. 283–294, 1997.

[12] J. Nocedal and S. J. Wright, *Numerical Optimization*, ser. Springer Series in Operations Research. Springer, 1999.

[13] V. Klumpp, F. Sawo, U. D. Hanebeck, and D. Fränken, "The Sliced Gaussian Mixture Filter for Efficient Nonlinear Estimation," in *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, Jul. 2008.