

Direct Fusion of Dirac Mixture Densities using an Efficient Approximation in Joint State Space

Vesa Klumpp and Uwe D. Hanebeck

Abstract—In this paper, we present a direct fusion algorithm for processing the combination of two Dirac mixture densities. The proposed approach allows the multiplication of two Dirac mixture densities without requiring identical support and thus enables the fusion of two independently generated sample sets. The resulting posterior Dirac mixture density is an approximation of the true continuous density that would result from the processing of the underlying true continuous density functions. This procedure is based on a suboptimal greedy approximation of the joint state space by means of a Dirac mixture that iteratively increases the resolution of the fusion result while considering only the relevant regions in the joint state space, where the fusion constraint holds.

NOTATION

\underline{x}	State vector $\in \mathbb{R}^N$
$\mathbf{x}, \underline{\mathbf{x}}$	Random variable and random vector
$\delta(x)$	Dirac delta function
$H(x)$	Heaviside step function
$\delta(\underline{x})$	Multi-dimensional Dirac delta function
$\tilde{f}(\underline{x})$	True density function (to be approximated)
$f(\underline{x})$	Approximate density function
$F(\underline{x})$	Cumulative distribution function
$\mathcal{N}(\cdot, m, \sigma)$	Gaussian density with mean m and standard deviation σ
$(\cdot)^p$	(Upper index p) Posterior quantity
$\underline{x} \downarrow Z$	Projection of \underline{x} onto subspace Z

I. INTRODUCTION

In many technical applications, the fusion of information given as probability densities is an important issue. An example of great interest lies in the area of processing information by means of sensor networks, which are used in observation of distributed phenomena [1]. Here, an important problem is the exchange of information between single sensor nodes [2], [3]. The computational task of the sensor nodes is to perform the data fusion in a decentralized manner where many nodes in the network perform the task, i.e., without a central processing node.

Unfortunately, in many such applications, the processing of the true continuous densities is either too complex or impossible. In order to overcome this problem, the density functions are approximated by parameterized functions that are more suitable for further processing. Well-known methods for estimation purposes are the Kalman filter [4] for linear systems and extensions of it, like the Extended Kalman

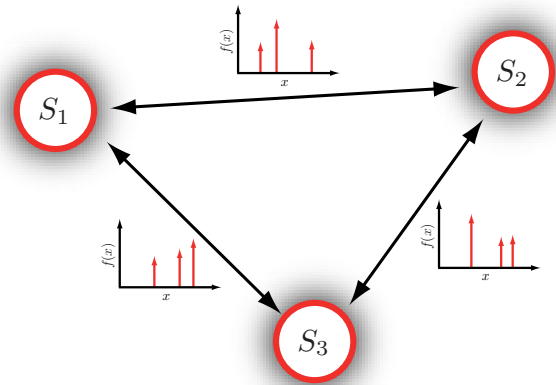


Fig. 1. Application example: Sensor network. Exchange of information in form of probability density functions given as Dirac mixture densities.

Filter [5] or Unscented Kalman Filter [6]. Unfortunately, the resulting Gaussian densities are not capable of representing arbitrary density functions that may appear in estimation problems of arbitrary nonlinear systems. Other density representations include Gaussian mixture densities [7], grid-based approaches [8], simple moments of probability density functions [9], exponential densities [3], Fourier series [10], [11], the representation by means of sample sets [12], or Dirac mixture densities [13], which are capable of representing more general densities.

The advantages of these parameterizations, like fast and simple processing, are at the expense of the versatility of arbitrary processing of these densities. Often, fixed processing steps are defined, which only solve some special problems, like the Bayesian filter step or prediction step in a special manner. Different mappings of density functions, like the computation of the point wise maximum or average, are often difficult to handle or intractable.

In this work, a special case of processing Dirac mixture densities on continuous domains is considered. The Dirac mixture representation, or sample set representation, is useful for a number of applications. It allows simple and fast processing, even for nonlinear systems, and the approximation of arbitrary probability density functions. A main problem of this representation is direct point wise multiplication, or the direct Bayesian inference of two Dirac mixture densities. Usually, the result of the multiplication of these density representations is not well defined as the two densities do not possess the same support. Here, the Dirac mixture densities are assumed to represent underlying true continuous densities, which is then exploited by approximating the

corresponding joint densities around the fusion constraint. The Dirac mixture resulting from the direct fusion, i.e., by evaluating the approximation of the joint density at the fusion constraint, is then an approximation of the continuous density that would result from fusing the underlying true continuous densities.

The paper is structured as follows. In the next Section, related works in the area of handling Dirac mixture functions are briefly described. Section III gives an introduction to Dirac mixture densities and the problem solved within this work. A general approach for handling this kind of problem by considering the joint state space for the processing function is shown in Section IV. In Section V the approximation of the joint state space by means of Dirac mixture that can be transformed into the posterior density is described in detail. In Section VI, the performance of the proposed procedure is depicted in different examples.

II. RELATED WORK

For approximating continuous density functions by means of Dirac mixture densities, different approaches exist. The probably best known and in many cases easy to implement algorithm is Monte Carlo sampling, applied in the particle filter [12]. A drawback of this sampling algorithm is the independent treatment of each sample, the non-determinism of the representation, and thus, the need for a large number of sample points. These problems are partially treated by deterministic quasi Monte Carlo sampling techniques [14].

Further improvements include the systematic approximation by minimizing a certain distance measure in cumulative distribution space. For that purpose, two procedures have been introduced: batch approaches and sequential approaches. Batch approaches calculate the optimal positions and weights of all Dirac mixture components simultaneously [13], [15], whereas sequential approaches are iterative algorithms that improve the approximation quality gradually by inserting single Dirac components at specific locations [16], [17]. This allows adjusting the approximation online, until the desired quality is obtained, or until the algorithm runtime is consumed, which enables the application of the given algorithm to be used in realtime scenarios.

The main problem addressed in this paper is Bayesian fusion of two sample sets given as Dirac mixture densities. This problem has been considered by a few authors only and has not been satisfactorily solved so far. Previous work focused on the reconstruction of the underlying continuous density from which the discrete sample sets are generated. In the field of Bayesian analysis, the problem of estimating parameters of continuous densities from samples is treated in [18]. Problems within this framework arise from assumptions on the underlying continuous density, like the kind of density class or the number of components and parameters used, which often do not comply with the true continuous density. Other approaches include the use of Gaussian mixture models and the estimation of their parameters by means of expectation maximization [19] and the use of Gaussian kernels that are placed onto every Dirac component. The choice of the

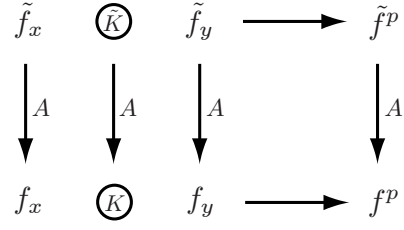


Fig. 2. Processing of two true densities \tilde{f}_x and \tilde{f}_y . The approximations A of the true densities and the processing step \tilde{K} give f_x , f_y , and K , which result in the posterior density f^P that is an approximation of the resulting true density f^P .

variance of the Gaussian kernel is crucial, which serves as a smoothing function of the discrete Dirac mixture density. In this paper, the fusion of two Dirac mixture densities will be performed directly without resorting to an intermediate smoothed representation.

III. PROBLEM FORMULATION

First, some general definitions about Dirac delta functions δ and their corresponding cumulative distribution, the Heaviside distribution function H , are made.

Definition III.1 The one-dimensional Dirac delta function is defined by

$$\delta(x - \xi) = \begin{cases} \text{undefined,} & x = \xi \\ 0, & \text{elsewhere} \end{cases}$$

with the property $\int_{-\infty}^{\infty} \delta(x - \xi) dx = 1$. The cumulative distribution function of a Dirac delta function is the Heaviside function

$$H(x - \xi) = \begin{cases} 0, & x < \xi \\ \frac{1}{2}, & x = \xi \\ 1, & x > \xi \end{cases}.$$

The N dimensional Dirac delta function is defined as the product of N one-dimensional delta functions

$$\delta(\underline{x} - \underline{\xi}) = \prod_{i=1}^N \delta(x_i - \xi_i) \quad \text{with } \underline{x}, \underline{\xi} \in \mathbb{R}^N.$$

Now, we consider two random vectors \underline{x} and \underline{y} on continuous domains $\underline{x} \in \mathbb{R}^N$ and $\underline{y} \in \mathbb{R}^N$ characterized by two Dirac mixture densities given by

$$\begin{aligned} f_x(\underline{x}) &= \sum_{i=1}^{L_x} w_i^x \delta(\underline{x} - \underline{\xi}_i^x), \\ f_y(\underline{y}) &= \sum_{i=1}^{L_y} w_i^y \delta(\underline{y} - \underline{\xi}_i^y), \end{aligned} \quad (1)$$

whereas the corresponding weights w_i^x and w_i^y sum up to 1 for each Dirac mixture and are all positive. The two mixtures generally possess a different number of components and do not have the same support.

Our goal is to perform a Bayesian fusion of the two random vectors \underline{x} and \underline{y} , i.e., multiplication with a subsequent renormalization, which at first look is not possible as the

multiplication of two Dirac mixtures with different support is not well-defined.

Here, we assume the Dirac mixture densities to be approximations of underlying true continuous densities \tilde{f}_x and \tilde{f}_y , see Fig. 2, that are unknown. Hence, a modified Bayesian fusion mechanism denoted by K has to be found that approximates the true fusion mechanism \tilde{K} in such a way that the *direct* fusion of the two Dirac mixtures results in a new Dirac mixture, which is a close approximation of the true continuous density \tilde{f}^p that would result from fusing the underlying true densities \tilde{f}_x and \tilde{f}_y .

IV. GENERAL APPROACH IN JOINT STATE SPACE

In this section, the general approach for the processing of two Dirac mixture densities in the joint state space is described.

Two N -dimensional random vectors \underline{x} and \underline{y} are described by their probability density functions f_x and f_y . The joint density is given as the $2N$ -dimensional Dirac mixture density

$$f(\underline{x}, \underline{y}) = \sum_{i=1}^{L_{xy}} w_i^{xy} \delta \left(\begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} - \underline{\xi}_i^{xy} \right) .$$

with weights w_i^{xy} and $2N$ -dimensional component locations $\underline{\xi}_i^{xy}$. Note, that in the case of independent random variables, the joint density function $f(\underline{x}, \underline{y})$ can be calculated according to the product $f(\underline{x}, \underline{y}) = f_x(\underline{x}) \cdot f_y(\underline{y})$.

In order to obtain the resulting posterior density $f^p(\underline{x})$, the joint density has to be evaluated according to the fusion constraint $\underline{y} = \underline{x}$, i.e., $f^p(\underline{x}) = f(\underline{x}, \underline{x})$. The problem is that this result is usually zero, because there is no mutual support of the single densities and thus the Dirac components $\underline{\xi}_i^{xy}$ in joint state space do not match to the fusion constraint.

In this paper we propose the approach of approximating the given Dirac mixture density $f(\underline{x}, \underline{y})$ by a representation that is well defined over the fusion constraint and can be converted into the desired posterior density $f^p(\underline{x})$ in the special case of combining two random vectors according to the constraint $\underline{y} = \underline{x}$. Thus, an appropriate representation of the joint density in the region of interest is needed without any knowledge of the underlying true densities \tilde{f}_x and \tilde{f}_y .

V. DIRAC APPROXIMATION

A. Approximation of Joint Density

The approximation approach is a special case of Dirac approximation of arbitrary multi-dimensional density functions [17]. Here, only Dirac mixture densities are considered.

Given are two Dirac mixture densities $f_x(\underline{x})$ and $f_y(\underline{y})$. The algorithm for the approximation of $2N$ -dimensional joint densities $\tilde{f}(\underline{x}, \underline{y})$ by means of Dirac mixture components is based on a greedy algorithm that improves the approximation quality at specific locations of interest. These regions have two things in common: the approximation quality over these regions is low, i.e., the approximation can be improved by inserting Dirac components into the approximation inside of these regions, and they comply with the constraint $\underline{y} = \underline{x}$ to a certain degree.

In this algorithm, the joint state space is divided into axis-aligned rectangular regions. Within each region, one Dirac component approximates the probability mass of the true density function.

Definition V.1 An axis-aligned rectangular region is defined by two vectors \underline{l} and \underline{u} . A point \underline{x} is inside of the region, if $\underline{l} \leq \underline{x} \leq \underline{u}$ holds. The region is referred to as $[\underline{l}, \underline{u}]$.

Remark V.1 Note, that for two vectors $\underline{x}, \underline{y} \in \mathbb{R}^N$ with $\underline{x} = [x_1, x_2, \dots, x_N]^T$ and $\underline{y} = [y_1, y_2, \dots, y_N]^T$ the relation $\underline{x} \leq \underline{y}$ holds, if and only if for every element i the relation $x_i \leq y_i, i \in \{1, \dots, N\}$ holds.

One Dirac component approximates the true underlying joint density $\tilde{f}(x, y)$ with the same probability mass

$$w = \int_{\underline{l}}^{\underline{u}} f(\underline{x}) d\underline{x} . \quad (2)$$

In the case of a Dirac mixture density $\tilde{f}(\underline{x}, \underline{y})$, the integral (2) turns into a simple sum of weights of the components inside $[\underline{l}, \underline{u}]$

$$w = \sum_{\{i: \underline{l} \leq \underline{\xi}_i \leq \underline{u}\}} w_i .$$

The following Lemma treats the approximation of two $2N$ -dimensional Dirac components by one Dirac component. This Dirac interpolation is used for positioning Dirac components that approximate the specified axis-aligned region with multiple Dirac components inside.

Lemma V.1 The Dirac position $\underline{\xi}$ interpolating two Dirac components at locations $\underline{\xi}_a$ and $\underline{\xi}_b$ is

$$\underline{\xi} = \frac{\underline{\xi}_a + \underline{\xi}_b}{2} .$$

Remark V.2 The resulting Dirac position $\underline{\xi}$ is used to group the Dirac components and associated weights into two groups. As it will be seen later, this Dirac position defines the location of the splitting line that is used to reduce the considered region, and furthermore maximizes entropy. Note, that this position is independent of the Dirac weights.

The position of the Dirac component that approximates the given region $[\underline{l}, \underline{u}]$ is defined by all one-dimensional sub-marginals of $2N$ -dimensional joint state space over the considered region, as defined in Definition V.2. In Lemma V.2 the position is derived.

Definition V.2 The one-dimensional sub-marginal of dimension j (described by subspace Z_j) inside of the region $[\underline{l}, \underline{u}]$ is a Dirac mixture density

$$f_j(z_j) = \sum_{i=1}^n w_{j,i} \delta(z_j - \xi_{j,i}) ,$$

with $z_j = [x, y]^T \downarrow Z_j$ and $\xi_{j,i} = \underline{\xi}_i \downarrow Z_j$, if $\underline{l} \leq [x, y]^T \leq \underline{u}$ holds. The weight $w_{j,i}$ is equal to the weight w_i of the

corresponding Dirac component in joint state space, if for the Dirac position $\underline{l} \leq \underline{\xi} \leq \underline{u}$ holds. For Diracs outside of the region, the weight $w_{j,i}$ is zero. If multiple components share the same position in the sub-marginal, they are merged and their weights are cumulated.

Lemma V.2 *The Dirac position in joint state space regarding all one-dimensional sub-marginals of the region $[\underline{l}, \underline{u}]$ is constructed by identifying the two sets of Dirac components that approximately halve the probability mass. The limiting components $\underline{\xi}_a$ and $\underline{\xi}_b$ are defined by their indices*

$$a = \arg \min_k \left\{ \left| \left(\sum_{i=1}^k w_{j,i} \right) - \left(\sum_{i=k+1}^n w_{j,i} \right) \right| \right\} \text{ and}$$

$$b = a + 1 .$$

With these components, the Dirac position $\underline{\xi}$ can be determined by applying the procedure successively to every dimension and using Lemma V.1.

The overall approximation algorithm starts with an $2N$ -dimensional axis-aligned region covering all the given components of the Dirac mixture $f(\underline{x}, \underline{y})$. Then, the probability mass is calculated and the first approximating Dirac component is inserted. In the next step, the region is split into two new regions. In the following steps, the region that will be split is chosen and the splitting is performed on it. These two steps are repeated until no more regions can be split or the desired approximation quality is reached. The complexity of the algorithm depends linearly on the number of Dirac components L_x and L_y of $f(\underline{x})$ and $f(\underline{y})$ and the number of dimensions N . The algorithm is shown in Algorithm 1.

Algorithm 1 Approximation Algorithm

- 1: initialize with one Dirac component
 - 2: **while** splitting candidates left or approximation quality below threshold **do**
 - 3: choose next splitting region
 - 4: determine splitting direction
 - 5: perform splitting
 - 6: update new Dirac components
 - 7: **end while**
-

B. Region Selection

The selection of the splitting region is an essential part of the algorithm, because it determines the overall behavior of the approximation algorithm. The refinement of the approximation has to be made at locations, where it contributes to the resulting posterior density, and at those regions, where the approximation quality is poor. Here, the selection process is simple. The first requirement is that the constraint $\underline{y} = \underline{x}$ holds at least for one point \underline{p} inside of $[\underline{l}, \underline{u}]$:

$$\exists \underline{p} = \begin{bmatrix} p_x \\ p_y \end{bmatrix} : p_y = p_x \wedge \underline{l} \leq \underline{p} \leq \underline{u} .$$

The second requirement is that the splitting region has maximum probability mass w . Here, it is assumed that the

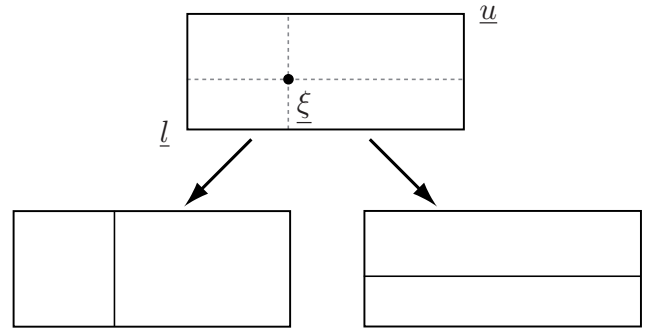


Fig. 3. Different splittings of a 2-dimensional region $[\underline{l}, \underline{u}]$ with the corresponding Dirac component $\underline{\xi}$. The given region can be split in two directions, resulting in two different sets of new sub-regions.

approximation quality, i.e., the error between the underlying true density function $\tilde{f}(\underline{x}, \underline{y})$ and the single Dirac component inside of a region, depends on the probability mass w . In this way it is guaranteed that important regions with high probability mass are favored.

For implementation, only these regions, in which the constraint $\underline{y} = \underline{x}$ holds, have to be kept, all others can be discarded. This keeps the required memory usage low and the considered region is subsequently narrowed towards the diagonal, which gives the desired result.

C. Splitting

The splitting introduced here improves the approximation quality of the joint density over the selected region. This is performed by splitting the given region into two new axis-aligned regions, whereas the union of them is the former region and the new probability masses sum up to the former probability mass.

The region is divided at the position $\underline{\xi}$ of the Dirac component along an axis, resulting in $2N$ potential splittings, e.g., in the case of a two-dimensional joint state space the splitting is performed in either x or y direction. The splitting direction is chosen in order to get quadratic regions, i.e., a regular approximation, and to minimize the considered region of the complete joint state space.

The first criterion defining the splitting dimension is constraint evaluation. A potential splitting candidate is chosen, if the plane $\underline{y} = \underline{x}$ intersects with only one of the new regions. In this case, the splitting only reduces the considered region in joint state space without increasing the number of Dirac components of the posterior approximation result. In the second criterion the region $[\underline{l}, \underline{u}]$ is split in the dimension i of it's maximum extent

$$i = \max_{i \in \{1, \dots, 2N\}} \{u_i - l_i\} .$$

Note, that regions containing only one Dirac component cannot be split further. This holds for the dimensions, whose sub-marginals contain one Dirac component only, too. Figure 4 shows the regions after 14 splitting steps. Within the yellow shaded regions the linear constraint is hold and these Dirac components are processed further.

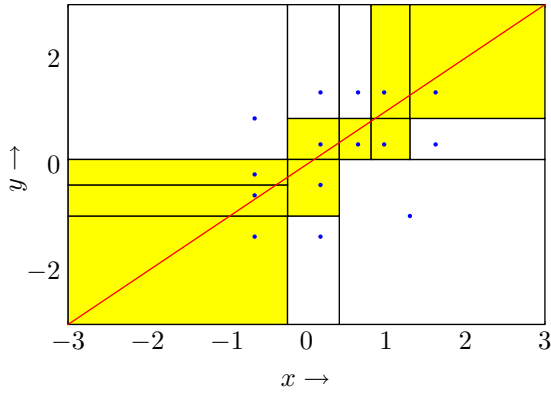


Fig. 4. Dirac approximation of 2-dimensional joint state space after 14 splittings. The constraint $y = x$ holds in the yellow regions. Computed Dirac positions are marked with a blue dot.

D. Posterior Density

In order to keep the equations simple, we consider a two-dimensional joint state space, i.e., $N = 1$, from here on. The approach for deriving the posterior Dirac mixture density

$$f^p(x) = \sum_{i=1}^{L^p} w_i^p \delta(x - \xi_i^p)$$

based on an approximation of the joint density will be explained now. Generally, two problems have to be solved, namely computing the weights and the positions of the individual Dirac mixture components. A straightforward approach is using the same weights w_i of the Dirac components in joint state space for the weights w_i^p of the posterior density function $f^p(x)$. Unfortunately, this leads to wrong results. The proposed algorithm approximated the joint state space, and thus the joint probability mass of a complete rectangular region, whereas for the resulting posterior density, only the density on the diagonal $y = x$, i.e., $f(x, x)$, shall be considered.

This discrepancy of probability representation of the region and the evaluation on a straight line can be handled in the following way, assuming uniform distribution of the probability mass w inside of $[l, u]$. This can be stated by having continuous and smooth true density functions $\tilde{f}_x(x)$ and $\tilde{f}_y(y)$, which result in a smooth and almost uniform distribution inside of $[l, u]$ in joint state space, assumed that the considered region is small enough, i.e., sufficient Dirac components are used for the approximation $f(x, y)$. Hence, the Dirac representation of the considered region can be transformed into a uniform density function

$$f_u(x, y) = f_u(z) = \begin{cases} \frac{w}{\prod_{i=1}^m u_i - l_i}, & l \leq z \leq u \\ 0, & \text{elsewhere} \end{cases} \quad (3)$$

Now, the probability density over the diagonal can be computed by evaluating a line integral over the region. For that purpose, the integration limits have to be defined, as in the following Lemma.

Lemma V.3 *The number of intersection points between the straight line $y = x$ and the boundary of the axis-aligned region $[l, u]$ is either zero or two. In case of intersection, the*

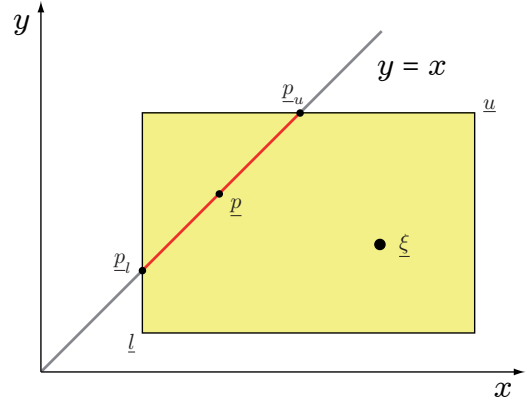


Fig. 5. The resulting Dirac component p is set at the center of the intersection of the considered region and the straight line $y = x$ and approximates the probability mass of the region $[l, u]$ on the diagonal.

points are at the locations

$$\begin{aligned} p_l &= [\max\{l\}, \max\{l\}]^T \in \mathbb{R}^2 \text{ and} \\ p_u &= [\min\{u\}, \min\{u\}]^T \in \mathbb{R}^2. \end{aligned}$$

Now, the probability mass w^p of the Dirac component of the posterior density function is

$$\begin{aligned} w^p &\propto \int_{p_l}^{p_u} \frac{w}{\prod_{i=1}^m u_i - l_i} dp \\ &= \sqrt{2} (\min\{u\} - \max\{l\}) \cdot \frac{w}{\prod_{i=1}^m u_i - l_i}. \end{aligned} \quad (4)$$

Note, that the probability mass w^p is only proportional to (4), because the complete approximation of joint state space is a normalized density. Thus, the diagonal $f(x, x)$ is never a normalized density function, however, the weights w^p can be normalized afterwards, after the processing of all posterior weights w_i^p .

The second step in computing the posterior density is the assignment of the Dirac position. Again, assuming uniform distribution over the considered region (3) leads to Theorem V.1, which states the optimal Dirac position \underline{p} . The geometrical layout of the region $[l, u]$, the straight line $y = x$, the intersection points p_l, p_u , and the resulting Dirac position \underline{p} are shown in Figure 5.

Theorem V.1 *The Dirac position minimizing the Cramér-von Mises distance measure of the posterior probability distribution for a given region is*

$$\underline{p} = \frac{p_l + p_u}{2}.$$

PROOF. Let x_p, x_l , and x_u be the projections of p, p_l and p_u onto the one-dimensional X -subspace. Let the probability mass w be uniformly distributed over the region $[l, u]$. w^p is the probability mass over the line integral according to (4). The projection of the uniform density over the line integral onto X gives the density

$$f_u^p(x) = \begin{cases} \frac{w^p}{x_u - x_l}, & x_l \leq x \leq x_u \\ 0, & \text{elsewhere} \end{cases}.$$

The one-dimensional Cramér-von Mises distance of the posterior distributions can be calculated according to

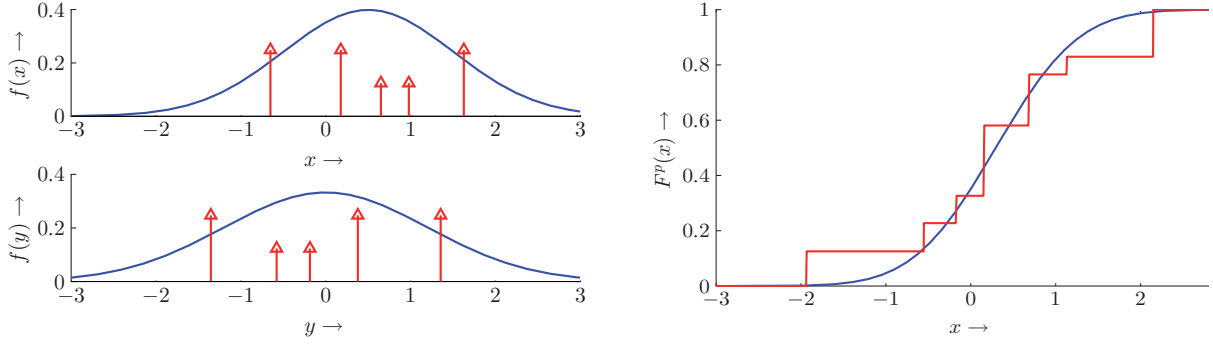


Fig. 6. Comparison of the processing of true density functions (blue) and their Dirac mixture approximations (red). The number of Dirac components used was both 5 for $f(x)$ and $f(y)$. The posterior density consists of 7 Dirac components.

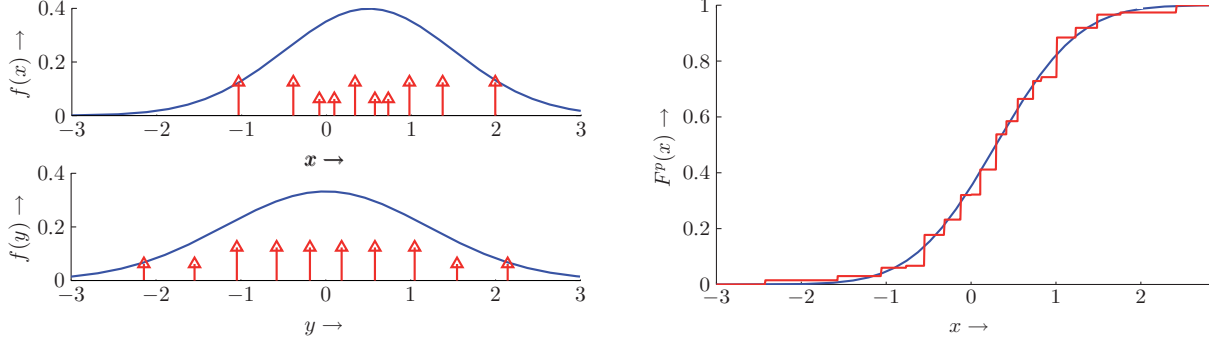


Fig. 7. Comparison of the processing of true density functions (blue) and their Dirac mixture approximations (red). The number of Dirac components used was both 10 for $f(x)$ and $f(y)$. The posterior density consists of 18 Dirac components.

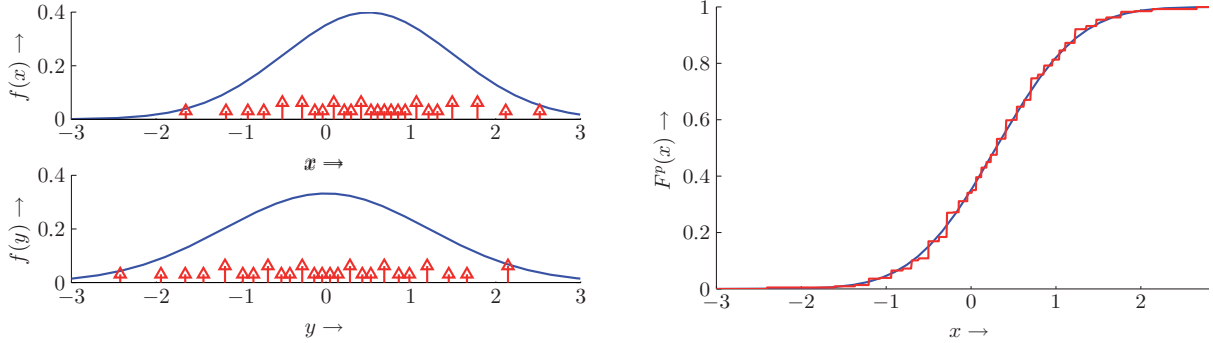


Fig. 8. Comparison of the processing of true density functions (blue) and their Dirac mixture approximations (red). The number of Dirac components used was both 25 for $f(x)$ and $f(y)$. The posterior density consists of 35 Dirac components.

$$\begin{aligned}
 D(x_p) &= \int_{-\infty}^{\infty} (F_u^p(x) - H(x - x_p))^2 dx \\
 &= \int_{x_l}^{x_u} \left(\frac{w^p \cdot (x - x_l)}{x_u - x_l} - H(x - x_p) \right)^2 dx .
 \end{aligned}$$

The root of the derivative

$$\frac{\partial}{\partial x_p} D(x_p) = \frac{(w^p)^2 (x_l + x_u - 2x_p)}{x_l - x_u}$$

is $\frac{x_l + x_u}{2}$, which leads with the condition $y = x$ directly to the desired result for p . \square

In the final step, the Dirac position p in joint state space is projected onto the X subspace, which gives the position ξ^p of the posterior Dirac mixture component. Note, that the optimality regarding the Cramér-von Mises distance can only be stated for one-dimensional problems.

VI. EXAMPLES

In this Section, examples are given for demonstrating the usefulness and performance of the proposed fusion approach

for two Dirac mixture densities. For that purpose, two known continuous densities are approximated by means of Dirac mixtures. Then, the result of fusing the Dirac mixtures is compared with the result of fusing the known continuous densities.

Example VI.1 This example shows the result of processing two Gaussian densities $\tilde{f}(x) = \mathcal{N}(x, 0.5, 1)$ and $\tilde{f}(y) = \mathcal{N}(y, 0, 1.2)$ directly and via their Dirac mixture approximations according to the algorithm described in [17]. Figure 6, Figure 7, and Figure 8 show the processed densities and the results for 5, 10 and 25 Dirac components used for both densities. In the upper left of each figure, the true continuous density $\tilde{f}(x)$ (blue) and its Dirac mixture representation (red) are shown. The lower left part displays the other density $\tilde{f}(y)$ and the corresponding Dirac mixture density. The processing result is shown on the right hand side. Cumulative distribution functions of the true distribution \tilde{F}^p (blue) and the approximation result (red) are shown. It is obvious, that the Heaviside mixture function representing the result of fusing the Dirac mixture approximations directly approaches the result of fusing the

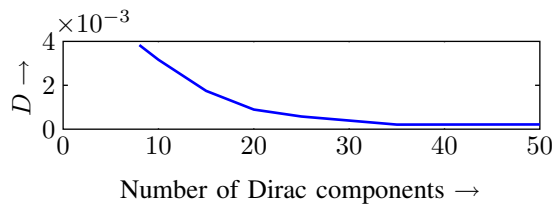


Fig. 9. Cramér-von Mises distance D between the approximated distribution $F^p(x)$ and the true distribution $\tilde{F}^p(x)$.

underlying true continuous distribution function with a growing number of components.

Example VI.2 In this example, the approximation quality is investigated for a different number of Dirac mixture components. The densities from Example VI.1 are approximated by two Dirac mixtures with the same number of components. Figure 9 displays the Cramér-von Mises distance D between the approximated posterior Distribution $F^p(x)$ and the true posterior distribution $\tilde{F}^p(x)$. The distance measure quickly decreases with a growing number of components.

VII. CONCLUSION AND FUTURE WORK

A new approach is proposed for the fusion of two Dirac mixture densities on continuous domains that are assumed to be approximations of underlying but unknown continuous densities. The Dirac mixtures are not required to possess the same support and are processed directly without prior smoothing, which results in a new Dirac mixture that approximates the (fictitious) result of fusing the true continuous densities. This is performed by approximating the corresponding joint density around the fusion constraint with a subsequent evaluation at the fusion constraint.

In a fusion mechanism, the new approach allows all the continuous densities to be represented by Dirac mixtures in an integrated manner and provides an efficient tool for their processing. Its performance has been demonstrated by simulations based on known underlying continuous densities.

In this paper, only the specific fusion constraint $x = y$ for the combination of two random vectors has been considered. Future work will be concerned with more general fusion constraints resulting from nonlinear measurement equations, which also includes the case of random vectors with different dimensions. As a result, including a measurement is different from just multiplying the component weights with the given likelihood function as performed in the particle filter measurement update and allows a systematic reapproximation of the update result.

In addition, the proposed new approach will be generalized to more general operations on two densities such as the point wise maximum or convolutions used for performing the prediction step in stochastic dynamic systems.

For dependent random quantities to be fused, we have to distinguish two cases. The first case, known dependencies, can easily be treated with the approach proposed in this paper by replacing the product of the two densities in the joint state space by the given joint density. The second case, unknown dependencies, is far more involved and the topic of current

research. It requires the reconstruction of the set of possible joint densities from their respective marginals.

REFERENCES

- [1] F. Sawo, K. Roberts, and U. D. Hanebeck, "Bayesian Estimation of Distributed Phenomena using Discretized Representations of Partial Differential Equations," in *Proceedings of the 3rd International Conference on Informatics in Control, Automation and Robotics (ICINCO 2006)*, Setúbal, Portugal, Aug. 2006, pp. 16–23.
- [2] B. S. Y. Rao, H. F. Durrant-Whyte, and J. A. Sheen, "A Fully Decentralized Multi-Sensor System for Tracking and Surveillance," *The International Journal of Robotic Research*, vol. 12, no. 1, pp. 20–44, 1993.
- [3] B. Tonkes and A. D. Blair, "Decentralized Data Fusion with Exponentials of Polynomials," in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, 2007.
- [4] R. E. Kalman, "A new Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME, Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.
- [5] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. McGraw-Hill, 2002.
- [6] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," in *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, Orlando, FL*, 1997.
- [7] D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian Estimation using Gaussian Sum Approximations," *IEEE Transactions on Automatic Control*, vol. 17, no. 4, pp. 439–448, Aug. 1972.
- [8] M. Šimandl, J. Kráľovec, and T. Söderström, "Anticipative Grid Design in Point-Mass Approach to Nonlinear State Estimation," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 699–702, 2002.
- [9] F. Carravetta, A. Germani, and M. Raimondi, "Polynomial Filtering of Discrete-Time Stochastic Linear Systems with Multiplicative State Noise," *IEEE Transactions on Automatic Control*, vol. 42, no. 8, pp. 1106–1126, 1997.
- [10] D. Brunn, F. Sawo, and U. D. Hanebeck, "Nonlinear Multidimensional Bayesian Estimation with Fourier Densities," in *Proceedings of the 2006 IEEE Conference on Decision and Control (CDC 2006)*, San Diego, California, Dec. 2006, pp. 1303–1308.
- [11] —, "Efficient Nonlinear Bayesian Estimation based on Fourier Densities," in *Proceedings of the 2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006)*, Heidelberg, Germany, Sept. 2006, pp. 312–322.
- [12] A. Doucet, N. D. Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, ser. Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001.
- [13] O. C. Schrempf, D. Brunn, and U. D. Hanebeck, "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, Dec. 2006.
- [14] G. Niederreiter, "Quasi-Monte Carlo methods and pseudo-random numbers," *Bulletin of the American Mathematical Society*, vol. 84, no. 6, pp. 957–1041, 1978.
- [15] O. C. Schrempf, D. Brunn, and U. D. Hanebeck, "Dirac Mixture Density Approximation Based on Minimization of the Weighted Cramér-von Mises Distance," in *Proceedings of the 2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006)*, Heidelberg, Germany, Sept. 2006, pp. 512–517.
- [16] 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.
- [17] V. Klumpp and U. D. Hanebeck, "Dirac Mixture Trees for Fast Sub-optimal 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.
- [18] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. Chapman & Hall/CRC, 1995.
- [19] X. Sheng, Y.-H. Hu, and P. Ramanathan, "Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network," in *Fourth International Symposium on Information Processing in Sensor Networks (IPSN)*, 2005.