Dirac Mixture Approximation of Multivariate Gaussian Densities

Uwe D. Hanebeck*, Marco F. Huber, and Vesa Klumpp

Abstract— For the optimal approximation of multivariate Gaussian densities by means of Dirac mixtures, i.e., by means of a sum of weighted Dirac distributions on a continuous domain, a novel systematic method is introduced. The parameters of this approximate density are calculated by minimizing a global distance measure, a generalization of the well–known Cramér–von Mises distance to the multivariate case. This generalization is obtained by defining an alternative to the classical cumulative distribution, the Localized Cumulative Distribution (LCD). In contrast to the cumulative distribution, the LCD is unique and symmetric even in the multivariate case. The resulting deterministic approximation of Gaussian densities by means of discrete samples provides the basis for new types of Gaussian filters for estimating the state of nonlinear dynamic systems from noisy measurements.

I. INTRODUCTION

Estimating the state of a nonlinear dynamic system from noisy observations is a difficult task. Even calculating a close approximation of the true posterior density for the prediction step or the filter step is very demanding from both a theoretical and a computational point of view.

However, for weakly nonlinear systems and unimodal posterior densities, a Gaussian approximation of the posterior density might be sufficient, which can be obtained by either a system approximation or a suitable approximation of the prior density. A system approximation is, for example, performed by the Extended Kalman Filter (EKF) by linearizing the system nonlinearities around the current estimate with a subsequent application of the Kalman filter. This, however, is of limited usefulness as the local linearization typically does neither sufficiently capture the global characteristics of the considered system nor the uncertainty of the prior density.

On the other hand, various approaches have been proposed for approximating the prior Gaussian density by sample points that are either selected randomly or carefully selected according to some optimality criterion [1], [2], [3]. These prior sample points are then propagated through the system and the measurement equation characterizing the given nonlinear system. Finally, based on the resulting posterior sample points, a suitable Gaussian approximation of the true (non–Gaussian) posterior density is calculated by means of, for example, moment matching. This Gaussian density can then be used recursively as a prior density for the next processing step. This paper focuses on the optimal approximation of prior Gaussian densities.

U. D. Hanebeck, M. F. Huber, and V. Klumpp are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics, Universität Karlsruhe (TH), Germany (e-mail: uwe.hanebeck@ieee.org, marco.huber@ieee.org, vesa.klumpp@ieee.org).

A Gaussian density with mean $\underline{m}$ and covariance matrix $\mathbf{\Sigma}$ according to

$$\tilde{f}(\underline{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{\Sigma}|} \exp\left(-\frac{1}{2}(\underline{x}-\underline{m})^T \mathbf{\Sigma}^{-1}(\underline{x}-\underline{m})\right) \tag{1}$$

characterizing a random vector $\underline{x} \in \mathbb{R}^N$ is approximated by a Dirac mixture density with L components given by

$$f(\underline{x}) = \sum_{i=1}^{L} w_i \cdot \delta(\underline{x}-\underline{x}_i) \ , \tag{2}$$

with positive weights, i.e., $w_i > 0$ for $i = 1, \ldots, L$, and locations $\underline{x}_i = \left[x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(N)}\right]^T$ for $i = 1, \ldots, L$.

Available approximation methods consider a random vector $\underline{x}_s$ distributed according to a standard Gaussian density given by

$$\tilde{f}(\underline{x}_s) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\frac{\left(x_s^{(k)}\right)^2}{\sigma^2}\right) \tag{3}$$

with zero mean and equal standard deviations σ in every dimension. The original random vector is recovered by the transformation

$$\underline{x} = \mathbf{S}\,\underline{x}_s + \underline{m} \ , \tag{4}$$

where $\mathbf{S}$ is some matrix square root of the covariance matrix $\mathbf{\Sigma}$ with $\mathbf{\Sigma} = \mathbf{S}\mathbf{S}^T$, e.g., the Cholesky decomposition of $\mathbf{\Sigma}$. After performing the approximation of the density of $\underline{x}_s$, the resulting Dirac mixture is transformed by using (4) to yield the desired approximation of the density of $\underline{x}$.

Several techniques for approximating a standard Gaussian density by a Dirac mixture are available in the literature, which can be divided into probabilistic and deterministic methods. Probabilistic methods draw samples from the given density, which can be implemented efficiently [4]. However, a lot of samples are typically required. Deterministic methods so far are limited to either placing components on a grid [5], which scales poorly for high–dimensional problems, or to placing components on the coordinate axes in order to avoid the curse of dimensionality. The simplest deterministic method, the so–called unscented transform [1], places two components along each coordinate axis and an optional component at the origin in such a way that the first two moments of the given density are matched. This technique has been generalized in [2] to match higher–order moments by placing more components along the coordinate axes. In [3], methods for the systematic Dirac mixture approximation of one–dimensional Gaussian densities by minimizing cumulative distance measures [6], [7] are exploited to again place the resulting components along the coordinate axes. Besides

considering shape information, this method also allows to match arbitrary moments.

In this paper, we propose a systematic method for multivariate approximation by minimizing an appropriate distance measure between the given Gaussian density $\tilde{f}(\underline{x})$ and its Dirac mixture approximation $f(\underline{x})$. For that purpose, it is not necessary to constrain the placement of the Dirac components in any way. In addition, it is possible to directly approximate Gaussian densities with different standard deviations in every dimension. Compared to the approximation of a standard Gaussian density with subsequent transformation, better approximation results are obtained, as the distance measure is not invariant to scaling. Hence, the scaling of the original density is left unchanged and we consider a random vector $\underline{x}_a$ characterized by a translated and rotated variant of the original density according to

$$\tilde{f}(\underline{x}_a) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\sigma^{(k)}} \exp\left(-\frac{1}{2} \frac{\left(x_a^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2}\right) \quad (5)$$

with zero mean and *different* standard deviations $\sigma^{(k)}$ in every dimension. With

$$\boldsymbol{\Sigma} = \mathbf{R}\,\text{diag}\left(\left[\sigma^{(1)}, \sigma^{(2)}, \ldots, \sigma^{(N)}\right]\right)\mathbf{R}^T$$

and $\mathbf{R}\mathbf{R}^T = \mathbf{I}$, the original random vector is recovered by

$$\underline{x} = \mathbf{R}\,\underline{x}_a + \underline{m}\ , \quad (6)$$

which also holds for obtaining the approximation of the density of the original random vector $\underline{x}$ from the approximation of the density of $\underline{x}_a$.

Standard density–based distance measures are not directly applicable to Dirac mixture densities, so distance measures based on cumulative distributions are typically employed. However, multivariate cumulative distributions exhibit the well-known problem of neither being unique nor symmetric [8]. Hence, an alternative to the classical cumulative distribution, the so called Localized Cumulative Distribution (LCD) is introduced in the next section. Based on this LCD, a generalization of the Cramér–von Mises distance, which is the squared integral difference between the LCD of the given density and the LCD of the approximate density is given in Section III. This new distance measure is used for determining the parameters of the approximate density in such a way that it is as close as possible to the given Gaussian density in Section IV. The new approach is discussed in Section V.

II. Localized Cumulative Distribution

In the scalar case, the classical cumulative distribution $F^c(m)$ is often employed for characterizing discrete random quantities or for comparing continuous and discrete random variables [6], [7]. However, when we consider the cumulative distribution $F^c(\underline{m})$ corresponding to a multivariate density function $g(\underline{x})$ defined by

$$F^c(\underline{m}) = \int_{-\infty}^{m^{(N)}} \cdots \int_{-\infty}^{m^{(1)}} g(\underline{x})\,\mathrm{d}x^{(1)}\ldots\mathrm{d}x^{(N)}\ ,$$

it is obvious that the integration performed over half–open infinite hyper–spaces makes $F^c(\underline{m})$ non–unique as different directions of integration are possible. In addition, due to the asymmetric integration, the resulting $F^c(\underline{m})$ is also not symmetric. While non–uniqueness and asymmetry do not pose a problem in the scalar case, the cumulative distribution is not well suited for comparing multivariate random quantities [8].

In contrast, the Localized Cumulative Distribution (LCD) considered in this paper is an alternative cumulative representation of a given random vector that performs an integration over symmetric kernels at all possible positions and for all kernels sizes.

Definition II.1 (Localized Cumulative Distribution) *Let $\underline{x}$ be a random vector with $\underline{x} \in \mathbb{R}^N$, which is characterized by an N–dimensional probability density function $g(\underline{x}) : \mathbb{R}^N \to \mathbb{R}_+$. The corresponding Localized Cumulative Distribution (LCD) is defined as*

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

with $\underline{b} \in \mathbb{R}_+^N$ and $F(.,.) : \Omega \to [0, 1]$, $\Omega \subset \mathbb{R}^N \times \mathbb{R}_+^N$. $K(\underline{x} - \underline{m}, \underline{b})$ is a suitable (symmetric & integrable) kernel located at position $\underline{m}$ and size characterized by $\underline{b}$ with $K(.,.) : \Omega \to [0, 1]$.

In this paper, we focus attention on separable kernels of the type

$$K(\underline{x} - \underline{m}, \underline{b}) = \prod_{k=1}^{N} K(x^{(k)} - m^{(k)}, b^{(k)})\ .$$

Furthermore, we consider kernels with equal width in every dimension, i.e., $b^{(k)} = b$ for $k = 1, \ldots, N$, which gives

$$K(\underline{x} - \underline{m}, b) = \prod_{i=k}^{N} K(x^{(k)} - m^{(k)}, b)\ .$$

In contrast to [8], where we considered rectangular kernels, the specific kernels used in this paper are of Gaussian type

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

Based on these Gaussian kernels, an N–dimensional Dirac component $\delta(\underline{x} - \hat{\underline{x}})$ at location $\hat{\underline{x}}$ corresponds to its LCD

$$\begin{aligned}
\Delta(\underline{m}, b) &= \int_{\mathbb{R}^N} \delta(\underline{x} - \hat{\underline{x}}) \cdot K(\underline{x} - \underline{m}, \underline{b})\,\mathrm{d}\underline{x} \\
&= \prod_{k=1}^{N} \exp\left(-\frac{1}{2}\frac{\left(\hat{x}^{(k)} - m^{(k)}\right)^2}{b^2}\right)\ .
\end{aligned}$$

As a result, the LCD of the Dirac mixture in (2) is given by

$$F(\underline{m}, b) = \sum_{i=1}^{L} w_i \prod_{k=1}^{N} \exp\left(-\frac{1}{2}\frac{\left(x_i^{(k)} - m^{(k)}\right)^2}{b^2}\right)\ . \quad (7)$$

The Gaussian density $\tilde{f}(\underline{x})$ according to (5) corresponds to an LCD

$$
\begin{aligned}
\tilde{F}(\underline{m}, b) &= \int_{\mathbb{R}^N} \tilde{f}(\underline{x})\, K(\underline{x} - \underline{m}, b)\, \mathrm{d}\underline{x} \\
&= \int_{-\infty}^{\infty} \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\, \sigma^{(k)}} \exp\left(-\frac{1}{2} \frac{\left(x^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2}\right) \\
&\qquad \prod_{k=1}^{N} \exp\left(-\frac{1}{2} \frac{\left(x^{(k)} - m^{(k)}\right)^2}{b^2}\right) \mathrm{d}\underline{x} \\
&= \prod_{k=1}^{N} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\, \sigma^{(k)}} \exp\left(-\frac{1}{2} \frac{\left(x^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2}\right) \\
&\qquad \exp\left(-\frac{1}{2} \frac{\left(x^{(k)} - m^{(k)}\right)^2}{b^2}\right) \mathrm{d}x^{(k)} \\
&= \prod_{k=1}^{N} \frac{b}{\sqrt{\left(\sigma^{(k)}\right)^2 + b^2}} \exp\left(-\frac{1}{2} \frac{\left(m^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2 + b^2}\right) \ .
\end{aligned}
$$
(8)

The LCD defined in this section is of versatile applicability and can be used as a suitable replacement in any expression involving the classical cumulative distribution. In the next section, it is used for generalizing the Cramér–von Mises distance.

III. MODIFIED CRAMÉR–VON MISES DISTANCE

Based on the definition of the Localized Cumulative Distribution (LCD), we will now present a modified version of the Cramér–von Mises distance [9] suitable for comparing multivariate random quantities, i.e., random vectors.

Definition III.1 (Modified Cramér–von Mises Distance)
The distance D between two densities $\tilde{f}(\underline{x}) : \mathbb{R}^N \to \mathbb{R}_+$ and $f(\underline{x}) : \mathbb{R}^N \to \mathbb{R}_+$ is given in terms of their corresponding LCDs $\tilde{F}(\underline{m}, b)$ and $F(\underline{m}, b)$ as

$$
D = \int_{\mathbb{R}_+} w(b) \int_{\mathbb{R}^N} \left(\tilde{F}(\underline{m}, b) - F(\underline{m}, b)\right)^2 \mathrm{d}\underline{m}\, \mathrm{d}b \ , \quad (9)
$$

where $w(b) : \mathbb{R}_+ \to [0, 1]$ is a suitable weighting function.

An intuitive interpretation of the distance measure is the following. For a fixed kernel at position $\underline{m}$ and size b, the probability masses of the continuous and the discrete density under the kernel are compared by calculating their LCDs. The distance measure now includes kernels of all sizes at all positions and compares the resulting probability masses. For that purpose, the integral of the square of the difference between the LCD of the true density $\tilde{f}(\underline{x})$ and the LCD of its approximation $f(\underline{x})$ is employed.

Theorem III.1 *For the LCDs of the Gaussian $\tilde{F}(\underline{m}, b)$ and the Dirac mixture $F(\underline{m}, b)$ and the weighting function*

$$
w(b) = \begin{cases} \frac{1}{b^{N-1}} & b \in [0, b_{max}] \\ 0 & elsewhere \end{cases} \ , \quad (10)
$$

the following expression for the distance D is obtained

$$
D = D_1 - 2D_2 + D_3 \ with \ D_i = \int_{\mathbb{R}_+} w(b)\, P_i\, \mathrm{d}b
$$

for $i = 1, 2, 3$,

$$
P_1 = \pi^{\frac{N}{2}} b^{2N} \prod_{k=1}^{N} \frac{1}{\sqrt{\left(\sigma^{(k)}\right)^2 + b^2}} \ ,
$$

$$
\begin{aligned}
P_2 = (2\pi)^{\frac{N}{2}} b^{2N} &\left(\prod_{k=1}^{N} \frac{1}{\sqrt{\left(\sigma^{(k)}\right)^2 + 2\,b^2}}\right) \\
&\sum_{i=1}^{L} w_i \exp\left(-\frac{1}{2} \sum_{k=1}^{N} \frac{\left(x_i^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2 + 2\,b^2}\right) \ ,
\end{aligned}
$$

$$
P_3 = \pi^{\frac{N}{2}} b^N \sum_{i=1}^{L} \sum_{j=1}^{L} w_i\, w_j\, \exp\left(-\frac{1}{2} \frac{T_{ij}}{2\,b^2}\right)
$$

and

$$
T_{ij} = \sum_{k=1}^{N} \left(x_i^{(k)} - x_j^{(k)}\right)^2 \ .
$$

PROOF. The proof is given in Appendix I. □

The formulas for the distance measure in Theorem III.1 require a *one–dimensional* numerical integration over the variable b. A closed–form expression for the third part of the distance measure, i.e., D_3, will be derived subsequently. Closed–form expressions for D_1, D_2 can be derived for the special case of a standard normal density (3), i.e., for equal variances in every dimension.

In contrast to the on–line solution proposed in this paper, the closed–form solutions for D_1, D_2 can be used for off–line approximation. This comprises an off–line step for calculating the optimal parameters of a Dirac mixture approximation of the standard normal density in (3) for a desired number of components L and an on–line step for transforming the obtained parameters back to the space of the given arbitrary density by means of (4). This procedure is of lower computational complexity for the on–line step compared to the full on–line approximation procedure proposed in this paper. However, as mentioned before, the approximation quality is worse, as the distance measure is not invariant to scaling.

A closed–form solution for the expression D_3 is given in the next theorem.

Theorem III.2 *For $w(b)$ according to (10) and large b_{max}, the following closed–form expression for D_3 is obtained*

$$
D_3 = \frac{\pi^{\frac{N}{2}}}{8} \sum_{i=1}^{L} \sum_{j=1}^{L} w_i w_j \left(4b_{max}^2 - C_b\, T_{ij} + \mathrm{xlog}(T_{ij})\right) \ ,
$$

with $\mathrm{xlog}(z) = z \cdot \log(z)$ and constants $C_b = \log(4\, b_{max}^2) - \Gamma$ and

$$
T_{ij} = \sum_{k=1}^{N} \left(x_i^{(k)} - x_j^{(k)}\right)^2 \ .
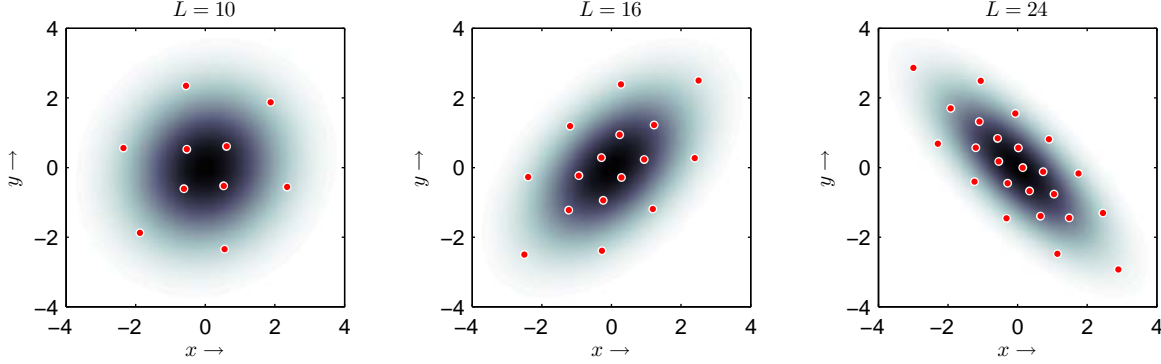$$

PROOF. The proof is given in Appendix II. □

Fig. 1. Approximations of two–dimensional Gaussian densities according to Example IV.1 for different numbers L of Dirac mixture components.

IV. APPROXIMATION

The desired parameters of the Dirac mixture in (2) approximating the given arbitrary multivariate Gaussian density in (1) are obtained by minimizing the distance measure in Theorem III.1. For that purpose, the Gaussian density (1) is first transformed to the zero–mean, axis–aligned (but unscaled) representation in (5). Then, the distance measure is minimized by means of some standard optimization routine. Here, we use the quasi–Newton method proposed in [10] for that purpose. Finally, the Dirac mixture obtained is transformed back to the original space by (6).

The optimization process reliably converges to the desired solution even when initialized with a set of randomly selected Dirac components. A formal derivation of conditions for uniqueness of the solution is currently investigated. However, intuitively the use of kernels of *all* sizes in the distance measure smoothes out unwanted local minima that exist for fixed kernel sizes and, hence, facilitates uniqueness.

For acceleration of the minimization procedure, the gradient of the distance measure is employed, which is now derived in analogy to the distance measure itself.

Theorem IV.1 *The gradient of the general distance measure in Theorem III.1 with respect to the locations of the Dirac components is given by*

$$G_\xi^{(\eta)} = \frac{\partial D}{\partial x_\xi^{(\eta)}} = G_\xi^{(\eta,1)} + G_\xi^{(\eta,2)}$$

with

$$G_\xi^{(\eta,1)} = 2\,(2\,\pi)^{\frac{N}{2}}\,w_\xi\,x_\xi^{(\eta)} \int_0^{b_{max}} \frac{b^{N+1}}{\left(\sigma^{(\eta)}\right)^2 + 2\,b^2}$$

$$\left(\prod_{k=1}^N \frac{1}{\sqrt{\left(\sigma^{(k)}\right)^2 + 2\,b^2}}\right) \exp\left(-\frac{1}{2}\sum_{k=1}^N \frac{\left(x_\xi^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2 + 2\,b^2}\right) \mathrm{d}b$$

and

$$G_\xi^{(\eta,2)} = -\pi^{\frac{N}{2}}\,w_\xi \sum_{i=1}^L w_i\left(x_\xi^{(\eta)} - x_i^{(\eta)}\right)$$

$$\int_0^{b_{max}} \frac{1}{b}\exp\left(-\frac{1}{2}\frac{T_{\xi i}}{2\,b^2}\right)\mathrm{d}b$$

for component index $\xi = 1, \ldots, L$ and dimension index $\eta = 1, \ldots, N$ with

$$T_{\xi i} = \sum_{k=1}^N \left(x_\xi^{(k)} - x_i^{(k)}\right)^2 \;.$$

PROOF. The proof is given in Appendix III. □

The gradient in Theorem IV.1 again requires a *one–dimensional* numerical integration over the variable b. However, a closed–form expression for $G_\xi^{(\eta,2)}$ can be found, which is given in the next theorem.

Theorem IV.2 *For large b_{max}, a closed–form expression for the second term of the gradient of the distance measure with respect to the locations of the Dirac components is given by*

$$G_\xi^{(\eta,2)} = \frac{\pi^{\frac{N}{2}}}{2} w_\xi \left\{ \sum_{i=1}^L w_i\left(x_\xi^{(\eta)} - x_i^{(\eta)}\right) \right.$$

$$\left. \log\left(\sum_{k=1}^N \left(x_\xi^{(k)} - x_i^{(k)}\right)^2\right) + C_b\left(\sum_{i=1}^L w_i x_i^{(\eta)} - x_\xi^{(\eta)}\right)\right\} \;,$$

for component index $\xi = 1, \ldots, L$ and dimension index $\eta = 1, \ldots, N$.

PROOF. The proof is given in Appendix IV. □

In many cases, it is desired to fulfill certain constraints, such as exactly matching certain moments of the given Gaussian or complying with certain restrictions in the state space like non–negative states. Arbitrary constraints are considered by solving an appropriate constrained optimization problem, where in this paper the mean and the covariance matrix of the given Gaussian are used as constraints. This can easily be achieved by Lagrangian relaxation of (9) or by including additive terms to (9) that penalize deviations from the true mean and covariance.

Example IV.1 *The approximation results for various two–dimensional Gaussian densities (w.l.o.g. zero mean) obtained with the proposed optimization procedure are shown in Figure 1. The following covariance matrices $\mathbf{\Sigma}$ are considered*

$$\begin{bmatrix} 2 & 0.2 \\ 0.2 & 2 \end{bmatrix}(left), \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}(mid), \begin{bmatrix} 2 & -1.5 \\ -1.5 & 2 \end{bmatrix}(right)$$

and the approximation is performed for equally weighted Dirac components with $L = 10, 16, 24$, respectively. It is obvious that the original continuous Gaussian density is systematically approximated by freely placed Dirac components without requiring any form of grid.
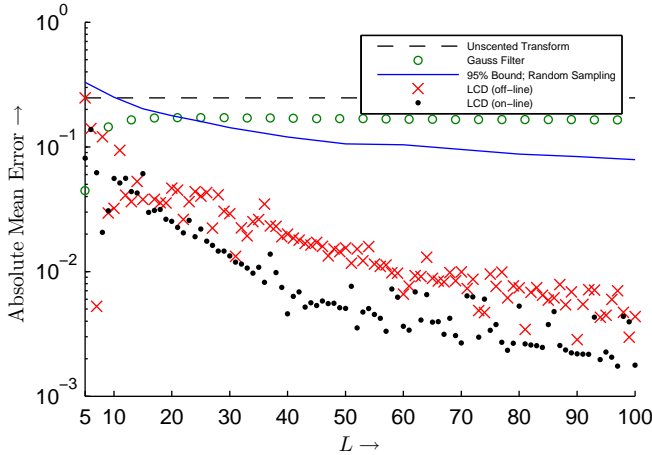
Fig. 2. Absolute errors of the mean $E(\boldsymbol{z})$ for various approximation algorithms in Example IV.2.

The next example demonstrates the performance of the proposed Dirac mixture approximation by a nonlinear transformation of a Gaussian random vector.

Example IV.2 *A random vector $\underline{\boldsymbol{x}} \in \mathbb{R}^2$ characterized by a zero–mean Gaussian density with covariance matrix $\Sigma = \mathrm{diag}([2, 0.2])$ is transformed to a scalar random variable $\boldsymbol{z}$ by*

$$\boldsymbol{z} = \sin(\boldsymbol{x}^{(1)})^2 + \cos(\boldsymbol{x}^{(2)})^2 \ .$$

Different types of Dirac mixture approximations of $\underline{\boldsymbol{x}}$ are employed to calculate the mean of $\boldsymbol{z}$:

- *Ground truth is obtained by using 10^7 random samples, which gives the mean $E^{true}(\boldsymbol{z}) = 0.6742$.*
- *Applying the unscented transform [1] yields the mean $E^{ut}(\boldsymbol{z}) = 0.9212$. The absolute error compared to the true mean is depicted as a dashed black line in Figure 2.*
- *The Gauss filter [3] for various numbers of components gives the absolute errors depicted by green circles in Figure 2.*
- *For every number of components L, 1000 different realizations of L random samples of $\underline{\boldsymbol{x}}$ are drawn, mapped to $\boldsymbol{z}$, and used for calculating an estimate of the mean of $\boldsymbol{z}$. Based on the resulting 1000 values for the absolute error between the mean estimates and the true mean, an error bound is drawn by the blue solid curve in Figure 2 that contains 95% of the absolute errors.*
- *The results of the Dirac mixture approximation based on an off–line approximation of a standard normal density discussed in Section III for various numbers of components L are given as red crosses in Figure 2.*
- *The on–line Dirac mixture approximation proposed in this paper gives the absolute errors shown as black dots in Figure 2 for various numbers of components L.*

In this example, the results of the UKF are far from the truth. The Gauss filter gives better results than the UKF but does not converge to the true result for an increasing number of components. Random sampling gives increasingly better results for a growing number of samples and eventually converges to the true results. However, convergence is slow and a large number of samples is required for ensuring satisfactory results. The Dirac mixture approximation based on an off–line approximation of a standard normal density gives satisfactory results even for small L and quickly approaches the true values for increasing L. The on–line Dirac mixture approximation proposed in this paper yields even better results and shows a faster convergence rate.

For different nonlinear mappings, the performance was found to be more or less unpredictable for the UKF, the Gauss filter, random sampling for a small number of components, and also for the Dirac

mixture approximation based on an off–line approximation of a standard normal density for a small number of components. The on–line Dirac approximation proposed in this paper reliably produced consistent results even for a small number of samples and quickly converged to the true results.

V. DISCUSSION

The Dirac mixture approximation proposed in this paper is used in several applications. One important application is Gaussian filtering of nonlinear stochastic dynamic systems as discussed in the introduction. Compared to standard approaches, performance has been shown to be significantly better as the full density is systematically approximated and an arbitrary number of components can be used.

In this paper, only the optimization of the positions of the Dirac mixture components with equal weights has been considered. Of course, non–equally weighted Dirac mixtures can be considered as well, where the weights are also optimized.

The proposed approximation of continuous densities by a discrete representation is completely different from available approaches like random sampling, Quasi Monte–Carlo, and vector quantization. In contrast to random sampling, no randomness is involved when calculating the proposed approximation. Instead, fully deterministic and repeatable results are obtained by solving an optimization problem. In addition, significantly less components are required thanks to the systematic placement, which is especially useful for high–dimensional systems.

Quasi Monte–Carlo methods employ low–discrepancy sequences [11], where a variety of methods for generating uniformly distributed sequences is available [12]. Unfortunately, for generating non–uniformly distributed low–discrepancy sequences only transformation methods are known that rely on the (typically unavailable) inverse distribution and suffer from the problem that the low–discrepancy property is in general not preserved for multi–dimensional random quantities [13]. The approach proposed in this paper directly approximates the given density and does not require any sort of distribution–based transformation method.

In vector quantization, the state space is tessellated into encoding regions that are represented by their centroids in such a way that the quantization error between the centroids and the corresponding points in the regions is minimized [14]. The joint optimization of both regions and centroids is inherently discontinuous, requires discrete decisions, and is prone to local minima. In the proposed method, no mapping of regions to discrete quantities is required. Instead, a global measure of the shape difference between the given continuous density and its discrete representation is minimized, which is inherently smooth. Hence, the desired parameters are obtained by applying standard optimization methods in a straightforward manner.

The paradigm of using LCDs can be exploited for calculating discrete approximations for arbitrary continuous densities, where the interesting special case of Gaussian mixture densities directly follows from the results of this paper. In addition, distance measures based on LCDs can be used for approximating discrete densities by continuous

densities (density estimation) or for performing a reduction of a given Dirac mixture density by representing it with fewer components.

Example code for Dirac mixture approximation of Gaussian densities can be found in the web at

<div align="center">

`http://isas.uka.de/material/`
`source-code/DM_Gauss_Approx` .

</div>

APPENDIX I
PROOF OF THEOREM III.1

The inner integral in the distance measure (9) is solved by exploiting the fact that the integration is over products of (unnormalized) Gaussian densities, which can be performed analytically. By inserting the LCD $\tilde{F}(\underline{m}, b)$, the first term can be written as

$$P_1 = \prod_{k=1}^{N} \int_{\mathbb{R}} \frac{b^2}{\left(\sigma^{(k)}\right)^2 + b^2} \exp\left(-\frac{1}{2}\frac{2\left(m^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2 + b^2}\right) dm^{(k)} .$$

Solving the N integrals gives

$$P_1 = \prod_{k=1}^{N} \sqrt{\pi} \frac{b^2}{\sqrt{\left(\sigma^{(k)}\right)^2 + b^2}} = \pi^{\frac{N}{2}} b^{2N} \prod_{k=1}^{N} \frac{1}{\sqrt{\left(\sigma^{(k)}\right)^2 + b^2}} .$$

Upon insertion of the LCDs $\tilde{F}(\underline{m}, b)$ and $F(\underline{m}, b)$, the second term is given by

$$P_2 = \int_{\mathbb{R}^N} \prod_{k=1}^{N} \frac{b}{\sqrt{\left(\sigma^{(k)}\right)^2 + b^2}} \exp\left(-\frac{1}{2}\frac{\left(m^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2 + b^2}\right)$$
$$\sum_{i=1}^{L} w_i \prod_{k=1}^{N} \exp\left(-\frac{1}{2}\frac{\left(x_i^{(k)} - m^{(k)}\right)^2}{b^2}\right) d\underline{m} .$$

Rearranging terms yields

$$P_2 = \sum_{i=1}^{L} w_i \prod_{k=1}^{N} \int_{\mathbb{R}} \frac{b}{\sqrt{\left(\sigma^{(k)}\right)^2 + b^2}} \exp\left(-\frac{1}{2}\frac{\left(m^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2 + b^2}\right)$$
$$\exp\left(-\frac{1}{2}\frac{\left(x_i^{(k)} - m^{(k)}\right)^2}{b^2}\right) dm^{(k)} ,$$

which after solving the N integrals over $m^{(k)}$, $k = 1, \ldots, N$ gives

$$P_2 = \sum_{i=1}^{L} w_i \prod_{k=1}^{N} \frac{\sqrt{2\pi}\, b^2}{\sqrt{\left(\sigma^{(k)}\right)^2 + 2 b^2}} \exp\left(-\frac{1}{2}\frac{\left(x_i^{(k)}\right)^2}{\left(\sigma^{(k)}\right)^2 + 2 b^2}\right).$$

Further simplification finally gives the desired result.

The third term P_3 is obtained by inserting the LCD $F(\underline{m}, b)$ as

$$P_3 = \int_{\mathbb{R}^N} \sum_{i=1}^{L}\sum_{j=1}^{L} w_i w_j \prod_{k=1}^{N} \exp\left(-\frac{1}{2}\frac{\left(x_i^{(k)} - m^{(k)}\right)^2}{b^2}\right)$$
$$\prod_{k=1}^{N} \exp\left(-\frac{1}{2}\frac{\left(x_j^{(k)} - m^{(k)}\right)^2}{b^2}\right) d\underline{m}.$$

Exchanging integration and summation gives

$$P_3 = \sum_{i=1}^{L}\sum_{j=1}^{L} w_i w_j \prod_{k=1}^{N} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\frac{\left(x_i^{(k)} - m^{(k)}\right)^2}{b^2}\right)$$
$$\exp\left(-\frac{1}{2}\frac{\left(x_j^{(k)} - m^{(k)}\right)^2}{b^2}\right) dm^{(k)}.$$

For further simplification, we use

$$\int_{\mathbb{R}} \exp\left(-\frac{1}{2}\frac{(z_i - m)^2}{b^2}\right) \exp\left(-\frac{1}{2}\frac{(z_j - m)^2}{b^2}\right) dm$$
$$= \sqrt{\pi}\, b \exp\left(-\frac{1}{2}\frac{(z_i - z_j)^2}{2 b^2}\right) , \tag{11}$$

which gives

$$P_3 = \pi^{\frac{N}{2}} b^N \sum_{i=1}^{L}\sum_{j=1}^{L} w_i w_j \prod_{k=1}^{N} \exp\left(-\frac{1}{2}\frac{\left(x_i^{(k)} - x_j^{(k)}\right)^2}{2 b^2}\right).$$

APPENDIX II
PROOF OF THEOREM III.2

For solving

$$D_3 = \int_{\mathbb{R}_+} \frac{1}{b^{N-1}} P_3 \, db ,$$

we define

$$\gamma(z) = \int_0^{b_{\max}} b \exp\left(-\frac{1}{2}\frac{z}{2 b^2}\right) db$$
$$= \frac{1}{8}\left\{4 b_{\max}^2 \exp\left(-\frac{1}{2}\frac{z}{2 b_{\max}^2}\right) + z \operatorname{Ei}\left(-\frac{1}{2}\frac{z}{2 b_{\max}^2}\right)\right\}$$

for $z > 0$, where $\operatorname{Ei}(.)$ is the exponential integral given by

$$\operatorname{Ei}(x) = \int_{-\infty}^{x} \frac{e^t}{t} dt .$$

Hence, D_3 is now given by

$$D_3 = \pi^{\frac{N}{2}} \sum_{i=1}^{L}\sum_{j=1}^{L} w_i\, w_j\, \gamma(T_{ij}) .$$

For small $x > 0$, the exponential integral can be approximated by

$$\operatorname{Ei}(-x) \approx \Gamma + \log(x) - x ,$$

where $\Gamma \approx 0.5772$ is the Euler gamma constant. Hence, the function $\gamma(z)$ can be approximated according to

$$\gamma(z) \approx \frac{1}{8}\left\{4 b_{\max}^2 \exp\left(-\frac{1}{2}\frac{z}{2 b_{\max}^2}\right)\right.$$
$$\left. + z\left(\Gamma + \log\left(\frac{1}{2}\frac{z}{2 b_{\max}^2}\right) - \frac{1}{2}\frac{z}{2 b_{\max}^2}\right)\right\}$$
$$\approx \frac{1}{8}\left\{4 b_{\max}^2 + z\left(\Gamma - \log(4 b_{\max}^2) + \log(z)\right)\right\}$$
$$= \frac{1}{8}\left\{4 b_{\max}^2 - C_b z + x\log(z)\right\}$$

with $C_b = \log\left(4 b_{\max}^2\right) - \Gamma$ and we obtain the desired result.

Taking the derivative of the distance measure in (9) with respect to a location $x_\xi^{(\eta)}$ gives

$$G_\xi^{(\eta)} = \frac{\partial D}{\partial x_\xi^{(\eta)}} = -2 \int_0^{b_{\max}} \frac{1}{b^{N-1}}$$
$$\int_{\mathbb{R}^N} \left(\tilde{F}(\underline{m}, b) - F(\underline{m}, b) \right) \frac{\partial F(\underline{m}, b)}{\partial x_\xi^{(\eta)}} \, \mathrm{d}\underline{m} \, \mathrm{d}b \ ,$$

with

$$\frac{\partial F(\underline{m}, b)}{\partial x_\xi^{(\eta)}} = -w_\xi \frac{x_\xi^{(\eta)} - m^{(\eta)}}{b^2}$$
$$\prod_{k=1}^N \exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(k)} - m^{(k)} \right)^2}{b^2} \right) \ .$$

In the following, we will treat the two parts of $G_\xi^{(\eta)}$ according to

$$G_\xi^{(\eta)} = G_\xi^{(\eta,1)} + G_\xi^{(\eta,2)}$$

separately. The first part is given by

$$G_\xi^{(\eta,1)} = -2 \int_0^{b_{\max}} \frac{1}{b^{N-1}} \int_{\mathbb{R}^N} \tilde{F}(\underline{m}, b) \frac{\partial F(\underline{m}, b)}{\partial x_\xi^{(\eta)}} \, \mathrm{d}\underline{m} \, \mathrm{d}b \ .$$

By using the expressions for $\tilde{F}(\underline{m}, b)$ and $\frac{\partial F(\underline{m},b)}{\partial x_\xi^{(\eta)}}$, we obtain

$$G_\xi^{(\eta,1)} = 2 w_\xi \int_0^{b_{\max}} \frac{1}{b^{N-1}} \int_{\mathbb{R}^N}$$
$$\prod_{k=1}^N \frac{b}{\sqrt{\left(\sigma^{(\eta)} \right)^2 + b^2}} \exp\left(-\frac{1}{2} \frac{\left(m^{(k)} \right)^2}{\left(\sigma^{(\eta)} \right)^2 + b^2} \right)$$
$$\frac{x_\xi^{(\eta)} - m^{(\eta)}}{b^2} \prod_{k=1}^N \exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(k)} - m^{(k)} \right)^2}{b^2} \right) \mathrm{d}\underline{m} \, \mathrm{d}b \ .$$

Combining the product terms gives

$$G_\xi^{(\eta,1)} = 2 w_\xi \int_0^{b_{\max}} \frac{1}{b^{N-1}} \int_{\mathbb{R}} \frac{x_\xi^{(\eta)} - m^{(\eta)}}{b^2}$$
$$\frac{b}{\sqrt{\left(\sigma^{(\eta)} \right)^2 + b^2}} \exp\left(-\frac{1}{2} \frac{\left(m^{(\eta)} \right)^2}{\left(\sigma^{(\eta)} \right)^2 + b^2} \right)$$
$$\exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(\eta)} - m^{(\eta)} \right)^2}{b^2} \right) \mathrm{d}m^{(\eta)}$$
$$\prod_{\substack{k=1 \\ k \neq \eta}}^N \int_{\mathbb{R}} \frac{b}{\sqrt{\left(\sigma^{(\eta)} \right)^2 + b^2}} \exp\left(-\frac{1}{2} \frac{\left(m^{(k)} \right)^2}{\left(\sigma^{(\eta)} \right)^2 + b^2} \right)$$
$$\exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(k)} - m^{(k)} \right)^2}{b^2} \right) \mathrm{d}m^{(k)} \mathrm{d}b \ .$$

For further simplification, we use

$$\int_{\mathbb{R}} \frac{b}{\sqrt{\sigma^2 + b^2}} \exp\left(-\frac{1}{2} \frac{m^2}{\sigma^2 + b^2} \right)$$
$$\exp\left(-\frac{1}{2} \frac{(x - m)^2}{b^2} \right) \mathrm{d}m$$
$$= \sqrt{2\pi} \frac{b^2}{\sqrt{\sigma^2 + 2 b^2}} \exp\left(-\frac{1}{2} \frac{x^2}{\sigma^2 + 2 b^2} \right)$$

and

$$\int_{\mathbb{R}} \frac{x - m}{b^2} \frac{b}{\sqrt{\sigma^2 + b^2}} \exp\left(-\frac{1}{2} \frac{m^2}{\sigma^2 + b^2} \right)$$
$$\exp\left(-\frac{1}{2} \frac{(x - m)^2}{b^2} \right) \mathrm{d}m$$
$$= \sqrt{2\pi} x \frac{b^2}{\left(\sigma^2 + 2 b^2 \right)^{\frac{3}{2}}} \exp\left(-\frac{1}{2} \frac{x^2}{\sigma^2 + 2 b^2} \right) \ ,$$

which gives

$$G_\xi^{(\eta,1)} = 2 \left(2\pi \right)^{\frac{N}{2}} w_\xi x_\xi^{(\eta)} \int_0^{b_{\max}} \frac{b^{N+1}}{\left(\sigma^{(\eta)} \right)^2 + 2 b^2}$$
$$\prod_{k=1}^N \frac{1}{\sqrt{\left(\sigma^{(k)} \right)^2 + 2 b^2}} \exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(k)} \right)^2}{\left(\sigma^{(k)} \right)^2 + 2 b^2} \right) \mathrm{d}b \ .$$

Simplifying the product over exponential functions gives the final result.

The second part is given by

$$G_\xi^{(\eta,2)} = 2 \int_0^{b_{\max}} \frac{1}{b^{N-1}} \int_{\mathbb{R}^N} F(\underline{m}, b) \frac{\partial F(\underline{m}, b)}{\partial x_\xi^{(\eta)}} \, \mathrm{d}\underline{m} \, \mathrm{d}b \ .$$

By using the expressions for $F(\underline{m}, b)$ and $\frac{\partial F(\underline{m},b)}{\partial x_\xi^{(\eta)}}$, we obtain

$$G_\xi^{(\eta,2)} = -2 w_\xi \int_0^{b_{\max}} \frac{1}{b^{N-1}} \int_{\mathbb{R}^N}$$
$$\sum_{i=1}^L w_i \prod_{k=1}^N \exp\left(-\frac{1}{2} \frac{\left(x_i^{(k)} - m^{(k)} \right)^2}{b^2} \right)$$
$$\frac{x_\xi^{(\eta)} - m^{(\eta)}}{b^2} \prod_{k=1}^N \exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(k)} - m^{(k)} \right)^2}{b^2} \right) \mathrm{d}\underline{m} \, \mathrm{d}b \ .$$

Combining the product terms gives

$$G_\xi^{(\eta,2)} = -2\,w_\xi \sum_{i=1}^{L} w_i \int_0^{b_{\max}} \frac{1}{b^{N-1}}$$

$$\int_{\mathbb{R}} \frac{x_\xi^{(\eta)} - m^{(\eta)}}{b^2} \exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(\eta)} - m^{(\eta)}\right)^2}{b^2}\right)$$

$$\exp\left(-\frac{1}{2} \frac{\left(x_i^{(\eta)} - m^{(\eta)}\right)^2}{b^2}\right) \mathrm{d}m^{(\eta)}$$

$$\prod_{\substack{k=1 \\ k \neq \eta}}^{N} \int_{\mathbb{R}} \exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(k)} - m^{(k)}\right)^2}{b^2}\right)$$

$$\exp\left(-\frac{1}{2} \frac{\left(x_i^{(k)} - m^{(k)}\right)^2}{b^2}\right) \mathrm{d}m^{(k)} \mathrm{d}b\,.$$

For further simplification, we use

$$\int_{\mathbb{R}} \frac{z_i - m}{b^2} \exp\left(-\frac{1}{2} \frac{(z_i - m)^2}{b^2}\right) \exp\left(-\frac{1}{2} \frac{(z_j - m)^2}{b^2}\right) \mathrm{d}m$$

$$= \sqrt{\pi} \frac{z_i - z_j}{2\,b} \exp\left(-\frac{1}{2} \frac{(z_i - z_j)^2}{2\,b^2}\right)$$

and (11), which leads to

$$G_\xi^{(\eta,2)} = -\pi^{\frac{N}{2}} w_\xi \sum_{i=1}^{L} w_i \left(x_\xi^{(\eta)} - x_i^{(\eta)}\right)$$

$$\int_0^{b_{\max}} \frac{1}{b} \prod_{k=1}^{N} \exp\left(-\frac{1}{2} \frac{\left(x_\xi^{(k)} - x_i^{(k)}\right)^2}{2\,b^2}\right) \mathrm{d}b$$

or equivalently

$$G_\xi^{(\eta,2)} = -\pi^{\frac{N}{2}} w_\xi \sum_{i=1}^{L} w_i \left(x_\xi^{(\eta)} - x_i^{(\eta)}\right)$$

$$\int_0^{b_{\max}} \frac{1}{b} \exp\left(-\frac{1}{2} \frac{T_{\xi i}}{2\,b^2}\right) \mathrm{d}b$$

for component index $\xi = 1, \ldots, L$ and dimension index $\eta = 1, \ldots, N$ with

$$T_{\xi i} = \sum_{k=1}^{N} \left(x_\xi^{(k)} - x_i^{(k)}\right)^2\,.$$

APPENDIX IV
PROOF OF THEOREM IV.2

By solving the integral

$$\int_0^{b_{\max}} \frac{1}{b} \exp\left(-\frac{1}{2} \frac{z}{2\,b^2}\right) \mathrm{d}b = -\frac{1}{2} \mathrm{Ei}\left(-\frac{1}{2} \frac{z}{2\,b_{\max}^2}\right)$$

for $z > 0$, we obtain

$$G_\xi^{(\eta,2)} = \frac{\pi^{\frac{N}{2}}}{2} w_\xi \sum_{i=1}^{L} w_i \left(x_\xi^{(\eta)} - x_i^{(\eta)}\right) \mathrm{Ei}\left(-\frac{1}{2} \frac{T_{\xi i}}{2\,b_{\max}^2}\right)$$

with

$$T_{\xi i} = \sum_{k=1}^{N} \left(x_\xi^{(k)} - x_i^{(k)}\right)^2$$

for component index $\xi = 1, \ldots, L$ and dimension index $\eta = 1, \ldots, N$. For large $b_{\max}$, we obtain

$$\mathrm{Ei}\left(-\frac{z}{4\,b_{\max}^2}\right) \approx \Gamma - \frac{z}{4\,b_{\max}^2} + \log\left(\frac{z}{4\,b_{\max}^2}\right)$$

$$\approx \Gamma - \log\left(4\,b_{\max}^2\right) + \log(z)$$

$$= -C_b + \log(z)$$

for $z > 0$. With

$$-C_b \sum_{i=1}^{L} w_i \left(x_\xi^{(\eta)} - x_i^{(\eta)}\right) = C_b \left(\sum_{i=1}^{L} w_i x_i^{(\eta)} - x_\xi^{(\eta)}\right)\,,$$

we obtain the desired result.

REFERENCES

[1] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," in *International Symposium on Aerospace/Defence Sensing, Simulation and Control*, 1997.

[2] D. Tenne and T. Singh, "The Higher Order Unscented Filter," in *Proceedings of the American Control Conference*, June 2003, pp. 2441–2446.

[3] 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)*, Seoul, Korea, Jul. 2008.

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

[5] K. Ito and K. Xiong, "Gaussian Filters for Nonlinear Filtering Problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 910–927, May 2000.

[6] 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.

[7] 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.

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

[9] D. D. Boos, "Minimum Distance Estimators for Location and Goodness of Fit," *Journal of the American Statistical Association*, vol. 76, no. 375, pp. 663–670, 1981.

[10] R. H. Byrd, P. Lu, and J. Nocedal, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific and Statistical Computing*, vol. 16, no. 5, pp. 1190–1208, 1995.

[11] D. Guo and X. Wang, "Quasi-Monte Carlo Filtering in Nonlinear Dynamic Systems," *IEEE Transactions on Signal Processing*, vol. 54, pp. 2087–2098, 2006.

[12] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, 1992.

[13] J. Hartinger and R. Kainhofer, "Non-uniform Low–discrepancy Sequence Generation and Integration of Singular Integrands," in *Monte Carlo and Quasi-Monte Carlo Methods 2004*, H. Niederreiter and D. Talay, Eds. Springer, 2006.

[14] G. Pagès and J. Printems, "Optimal Quadratic Quantization for Numerics: The Gaussian Case," *Monte Carlo Methods and Applications*, vol. 9, no. 2, pp. 135–166, 2003.