Progressive Gaussian Filtering Based on Dirac Mixture Approximations

Uwe D. Hanebeck, Jannik Steinbring
Intelligent Sensor-Actuator-Systems Laboratory (ISAS),
Institute for Anthropomatics,
Karlsruhe Institute of Technology (KIT), Germany.
Email: uwe.hanebeck@ieee.org, jannik.steinbring@student.kit.edu

***Abstract*—In this paper, we propose a progressive Gaussian filter, where the measurement information is continuously included into the given prior estimate (although we perform observations at discrete time steps). The key idea is to derive a system of ordinary first-order differential equations (ODE) that is used for continuously tracking the true non-Gaussian posterior by its best matching Gaussian approximation. Calculating the required moments of the true posterior is performed based on corresponding Dirac Mixture approximations. The performance of the new filter is evaluated in comparison with state-of-the-art filters by means of a canonical benchmark example, the discrete-time cubic sensor problem.**

I. Introduction

We consider state estimation in discrete-time stochastic nonlinear dynamic systems. Thanks to their simplicity, Gaussian filters, i.e., filters representing all state densities by Gaussians, are an attractive tool for solving this type of estimation problem. However, in general, their full estimation performance is not exploited due to additional assumptions and simplifications.

In this paper, we focus on Gaussian filters that operate by finding the best-matching Gaussian to the true posterior by means of an explicit shape optimization after every processing step. A filter of this type is called Gaussian-assumed density filter (GADF). During the filter step, GADFs automatically consider the actual measurement during the approximation.

Another commonly used type of Gaussian filter is the so called Linear Regression Kalman Filter (LRKF) [1]. LRKFs approximate the (non-Gaussian) joint density of state and measurement corresponding to the prior density, the noise density, and the given nonlinear measurement equation by a jointly Gaussian density, which does not consider the actual measurement and typically is only a very rough approximation of the true joint density. The actual measurement is then used for slicing the approximate joint density in order to obtain the posterior density. Due to this additional Gaussian assumption, performance is typically worse compared to GADFs not making this assumption.

Examples of LRKFs [2] are the Unscented Kalman Filter (UKF) [3] and its scaled version [4], its higher-order generalization [5], a generalization to an arbitrary number of deterministic samples placed along the coordinate axes [6], filters performing an analytic or semi-analytic calculation of the required moments [7], [8] based on a decomposition into parts that can be calculated in closed form or via a sample approximation [9], and filters based on numerical integration for calculating the required nonlinear moments of the prior Gaussian density [10].

Of course, GADFs are more complicated to implement in comparison to the LRKFs and there are various options for minimizing the shape deviation between the true posterior and its Gaussian approximation. One option is to employ moment matching, i.e., using the mean and covariance matrix of the true posterior as parameters for the desired Gaussian, as this is known to minimize the Kullback-Leibler distance between the two densities. Unfortunately, in the case of nonlinear measurement equations and the corresponding complicated Likelihood function, it is in general not a simple task to calculate mean and covariance matrix of the true posterior, as analytic solutions are a rare exception. In most cases, numerical integration over the true posterior, i.e., the product of the (Gaussian) prior density and the Likelihood function, is required, such as Monte-Carlo integration [11].

In this paper, we propose a progressive Bayesian procedure for Gaussian-assumed density filtering, where the measurement information is continuously included into the given prior estimate (although we perform observations at discrete time steps). The first progressive filtering procedure of this type has been developed in 2003 [12] for state estimation with a Gaussian Mixture representation in the scalar case, where a homotopy continuation approach was proposed for tracking the true posterior with its approximation minimizing a squared-integral distance measure. The multi-dimensional case was treated in [13]. A generalization of this method to various other distance measures is proposed in [14]. Besides state estimation, the progressive processing idea has been applied to

moment calculation [15] and Gaussian Mixture reduction [16]. A homotopy-based filtering method operating on a particle representation is given in [17].

This paper is based on the method in [13] that was developed for state estimation with a Gaussian Mixture representation. The key idea is to derive a system of ordinary first-order differential equations (ODEs) for continuously tracking the true non-Gaussian posterior by its best matching Gaussian approximation. In contrast to [13], where a squared-integral distance measure between the true and the approximate posterior was minimized, here we derive ODEs for directly tracking the mean and covariance of the true posterior density by its Gaussian approximation.

The new progressive estimation method allows for arbitrary noise structures, even for noise structures that cannot easily be treated by LRKFs such as multiplicative noise[1]. The required integrals are solved by employing a Dirac Mixture approximation [18] of the Gaussian posterior. Dirac Mixture approximations for nonlinear estimation and filtering have been proposed for the case of scalar continuous densities in [19], [20]. An algorithm for sequentially increasing the number of components is given in [21] and applied to recursive nonlinear prediction in [22]. Multi-dimensional Gaussian densities are treated in [18]. Of course, more complicated continuous densities can be handled with this approach such as Gaussian Mixture densities, which is outside the scope of this paper.

Progressive processing automatically places Dirac components solely in the interesting regions of the state space, i.e., the support of the true posterior. As a result, the new filtering method is fast, efficient, and robust. Its performance is evaluated in comparison with state-of-the-art filters by means of a canonical benchmark example, the discrete-time cubic sensor problem.

II. Problem Formulation

We consider the general problem of estimating the hidden state of a discrete-time stochastic nonlinear dynamic system based on noisy measurements, which consists of a prediction step (or time update) employing a system model for propagating the estimated state from time step to time step and a filter step (or measurement update) for including observations taken at a given time step into the state estimate. Here, the focus is on the filter step that is typically considered harder compared to the prediction step. The insights obtained for the filter step can, however, be used for the prediction step as well.

A generative measurement equation

$$\hat{\underline{y}} = \underline{h}(\underline{x}, \underline{v}) \tag{1}$$

is investigated, with state $\underline{x}$, a specific measurement $\hat{\underline{y}}$, and measurement noise $\underline{v}$ with corresponding noise density $f_v(\underline{v})$.

[1]Of course, the corresponding Likelihoood function has to be derived before applying the filter.

The special case of additive noise

$$\hat{\underline{y}} = \underline{h}(\underline{x}) + \underline{v} \tag{2}$$

is also of interest, as it usually simplifies matters.

We assume that the generative model can somehow be converted to a probabilistic model represented by the conditional density $f(\underline{y}|\underline{x})$. For a given specific measurement $\hat{\underline{y}}$, this conditional density is the so called Likelihood function[2] abbreviated as

$$f_L(\underline{x}) = f(\hat{\underline{y}}|\underline{x}) \ .$$

For the case of additive noise, the Likelihood function is given by

$$f_L(\underline{x}) = f_v(\hat{\underline{y}} - \underline{h}(\underline{x})) \ .$$

Other noise structures result in different Likelihood functions.

Gaussian Filters

For a Gaussian-assumed density filter, we have a Gaussian prior density $f^p(\underline{x})$ that undergoes a Bayesian filter step according to the following multiplication with the Likelihood function

$$\tilde{f}^e(\underline{x}) = f^p(\underline{x}) \cdot f_L(\underline{x}) \ ,$$

where the resulting true posterior density is denoted by $\tilde{f}^e(\underline{x})$. Please note that no normalization has been performed so that the posterior density does not necessarily integrate to one. The tilde is used to underline that this is the true density resulting from performing a single filter step.

Obviously, the true posterior density $\tilde{f}^e(\underline{x})$ in general is not Gaussian anymore. In order to enable recursive processing without increase in computational complexity, the true posterior has to be approximated by a Gaussian density for the next processing step. In this paper, this will be performed by moment matching, i.e., by calculating the Gaussian density with the same mean and covariance matrix as $\tilde{f}^e(\underline{x})$.

The key idea for calculating mean and covariance matrix of $\tilde{f}^e(\underline{x})$ is given in the next section.

III. Key Idea

Our goal is to calculate mean and covariance matrix of $\tilde{f}^e(\underline{x})$, which typically cannot be performed analytically. One option is to use numerical integration methods based on a discrete approximation of the prior density $f^p(\underline{x})$. This includes Monte-Carlo integration based on independent random samples drawn from $f^p(\underline{x})$ or replacing $f^p(\underline{x})$ by its deterministic Dirac Mixture approximation. However, the problem is that for narrow Likelihood functions, only a few discrete samples actually contribute to the integral and the remaining samples are wasted.

[2]The Likelihood function is not necessarily a valid density function. Although it is always non-negative, it does not necessarily integrate to one, i.e., it usually is not normalized or even cannot be normalized.

In this paper, the key idea is to gradually include the measurement information, so that intermediate Gaussian posteriors become available. These intermediate posteriors can then be discretized and used for numerical integration. As a result, samples can always be maintained in regions of high likelihood.

Gradual inclusion of the measurement information could be performed by decomposing the original Likelihood function into a finite product of (typically wider) Likelihood functions [23], [24]. This, of course, leads to the problem of determining the number of sub-Likelihoods used.

Here, we propose to employ a continuous inclusion of the measurement information, which is a specific form of homotopy continuation and gives two advantages. First, we can derive an elegant solution for tracking the parameters of the desired intermediate Gaussian posteriors in the form of a system of ordinary first-order differential equations. When solving this system on a digital computer, the optimal step sizes are automatically determined by the solver. Second, a continuous inclusion leads to a continuous change of the intermediate posteriors, which allows to use density discretization methods such as those in [20] that make use of homotopy continuation methods anyway.

We employ a Gaussian representation of the intermediate posteriors called f_c^e (as it is a continuous representation) with parameter vector $\underline{\eta}_c$. A discrete representation called f_d^e is maintained simultaneously. The two representations are tied in such a way that changing the shape of f_c^e directly influences the parameters of its discrete approximation f_d^e.

In the general multivariate case, the Dirac Mixture approximation of the Gaussian is calculated as described in [18]. For the scalar Gaussian densities used in the evaluation (Section VI), the explicit formulas from [19] are used for calculating Dirac Mixture approximations with equal weights.

IV. Progressive Filter Step

We will now derive the progressive filter step. The first step is to redefine the Likelihood function in such a way that a continuous execution of the filter step is achieved. This progressive Likelihood function is defined by

$$f_L(\underline{x}, \gamma) \ ,$$

where γ is an artificial time with $\gamma \in [0, 1]$. It is desired that

$$f_L(\underline{x}, \gamma) = \begin{cases} 1 & \gamma = 0 \\ f_L(\underline{x}) & \gamma = 1 \end{cases}$$

holds.

Several options for defining progressive Likelihood functions exist. This includes progressively modifying the given nonlinear mapping $\underline{h}(\underline{x})$ of the underlying generative model (1) or varying the noise variance as in [13]. Here, we use the exponentiation of the Likelihood function as used in [14], [17]. The modified Likelihood function is then given by

$$f_L(\underline{x}, \gamma) = [f_L(\underline{x})]^\gamma \ . \tag{3}$$

Remark IV.1 (Additive Noise) *For the special case of generative systems suffering from additive noise in (2), we now obtain the expression*[3]

$$f_L(\underline{x}, \gamma) = \exp\left\{ -\frac{1}{2} \gamma \ \left(\underline{y} - \underline{h}(\underline{x})\right)^T \mathbf{C}_v^{-1} \left(\underline{y} - \underline{h}(\underline{x})\right) \right\} \ ,$$

where scaling factors have been omitted to achieve $f_L(\underline{x}, \gamma = 0) = 1$.

In the following, the derivative of the Likelihood function with respect to γ is required. For the specific progression from (3), we obtain

$$\begin{aligned} \dot{f}_L(\underline{x}, \gamma) &= \frac{\partial f_L(\underline{x}, \gamma)}{\partial \gamma} \\ &= \frac{\partial f_L^\gamma(\underline{x})}{\partial \gamma} \\ &= f_L^\gamma(\underline{x}) \cdot \log\left(f_L(\underline{x})\right) \\ &= f_L(\underline{x}, \gamma) \cdot \log\left(f_L(\underline{x})\right) \ . \end{aligned}$$

Remark IV.2 (Additive Noise) *A further simplification is achieved by focusing on the case of additive noise. Taking the derivative of the Likelihood function now gives*

$$\dot{f}_L(\underline{x}, \gamma) = -\frac{1}{2} f_L(\underline{x}, \gamma) \left(\underline{y} - \underline{h}(\underline{x})\right)^T \mathbf{C}_v^{-1} \left(\underline{y} - \underline{h}(\underline{x})\right) \ .$$

The second step is the continuous execution of the filter step, now written with prior density $f^p(\underline{x})$ and an intermediate posterior density $\tilde{f}^e(\underline{x}, \gamma)$ depending on the artificial time γ introduced above

$$\tilde{f}^e(\underline{x}, \gamma) = f^p(\underline{x}) \cdot f_L(\underline{x}, \gamma)$$

for $\gamma \in [0, 1]$. For the final time[4] $\gamma = 1$, the progressive Likelihood reaches the original Likelihood and as a result, the original posterior is attained by the intermediate posterior, i.e., $\tilde{f}^e(\underline{x}, \gamma = 1) = \tilde{f}^e(\underline{x})$. On the other extreme, for the start time $\gamma = 0$, the intermediate posterior is desired to be identical to the prior density, i.e., we have $\tilde{f}^e(\underline{x}, \gamma = 0) = f^p(\underline{x})$.

Plugging in a continuous approximation $f_c^e(\underline{x}, \underline{\eta}_c(\gamma))$ of the true intermediate posterior $\tilde{f}^e(\underline{x}, \gamma)$ gives

$$f_c^e(\underline{x}, \underline{\eta}_c(\gamma)) \approx \tilde{f}^e(\underline{x}, \gamma) = f^p(\underline{x}) \cdot f_L(\underline{x}, \gamma) \tag{4}$$

[3]For the additive noise case, exponentiation is equivalent to progressive modification of the noise covariance matrix as proposed in [13] since we have

$$\gamma \left(\underline{y} - \underline{h}(\underline{x})\right)^T \mathbf{C}_v^{-1} \left(\underline{y} - \underline{h}(\underline{x})\right) = \left(\underline{y} - \underline{h}(\underline{x})\right)^T \left(\frac{1}{\gamma}\mathbf{C}_v\right)^{-1} \left(\underline{y} - \underline{h}(\underline{x})\right) \ .$$

[4]Of course, time here denotes the artificial time γ introduced above.

for $\gamma \in [0,1]$, where left-hand-side and right-hand-side now become equal only for $\gamma = 0$, but only approximately[5] equal for $\gamma > 0$.

V. Scalar Gaussian-Assumed Density Filter

This section is devoted to deriving specific formulas for a scalar Gaussian density for representing f_c^e given by

$$f_c^e(x, \underline{\eta}_c(\gamma)) = w(\gamma)\, \frac{1}{\sqrt{2\pi}\,\sigma(\gamma)} \exp\left(-\frac{1}{2}\frac{(x-m(\gamma))^2}{(\sigma(\gamma))^2}\right) \ ,$$

with

$$\underline{\eta}_c(\gamma) = [w(\gamma),\, m(\gamma),\, \sigma(\gamma)]^T \ .$$

An additional weighting factor $w(\gamma)$ has been introduced that allows tracking unnormalized true posteriors.

In order to find the best matching Gaussian approximation $f_c^e(x, \underline{\eta}_c(\gamma))$ in (4), we desire the first moments to be equal as

$$\int_{\mathbb{R}} \underline{m}\, f_c^e(x, \underline{\eta}_c(\gamma))\, dx = \int_{\mathbb{R}} \underline{m}\, f^p(x) \cdot f_L(x,\gamma)\, dx \ ,$$

with $\underline{m} = [1, x, x^2]^T$. Taking the derivative with respect to γ on both sides gives

$$\int_{\mathbb{R}} \underline{m}\, \frac{\partial f_c^e(x, \underline{\eta}_c(\gamma))}{\partial \gamma}\, dx = \int_{\mathbb{R}} \underline{m}\, f^p(x) \cdot \underbrace{\frac{\partial f_L(x,\gamma)}{\partial \gamma}}_{\dot{f}_L(x,\gamma)}\, dx \ ,$$

with

$$\frac{\partial f_c^e(x, \underline{\eta}_c(\gamma))}{\partial \gamma} = \underbrace{\frac{\partial f_c^e(x, \underline{\eta}_c(\gamma))}{\partial\, \underline{\eta}_c^T(\gamma)}}_{\underline{p}^T(x,\gamma)} \cdot \underbrace{\frac{\partial\, \underline{\eta}_c(\gamma)}{\partial \gamma}}_{\dot{\underline{\eta}}_c(\gamma)} \ .$$

We obtain

$$\underbrace{\int_{\mathbb{R}} \underline{m}\, \underline{p}^T(x,\gamma)\, dx}_{\mathbf{P}(\gamma)}\, \dot{\underline{\eta}}_c(\gamma) = \underbrace{\int_{\mathbb{R}} \underline{m}\, f^p(x) \cdot \dot{f}_L(x,\gamma)\, dx}_{\underline{b}(\gamma)} \ ,$$

which can be written as

$$\mathbf{P}(\gamma)\, \dot{\underline{\eta}}_c(\gamma) = \underline{b}(\gamma) \ . \tag{5}$$

[5]As long as working with the infinite-dimensional functional representation $\tilde{f}^e$ of the true density, the two sides in (4) are equal as $\tilde{f}^e(\underline{x}, \gamma)$ is capable of following changes in the right-hand-side exactly. On the other hand, for a finite-dimensional representation $f_c^e(\underline{x}, \underline{\eta}_c(\gamma))$, i.e., a density function depending on a finite-dimensional parameter vector $\underline{\eta}_c(\gamma)$, the left-hand-side cannot necessarily exactly follow the changes of the right-hand-side as the product on the right-hand-side typically is not of the same density type.

Closed-form Expressions for Matrix $\mathbf{P}(\gamma)$: The required derivatives of the continuous posterior f_c^e with respect to the density parameters are collected in the vector $\underline{p}(x,\gamma)$ given by

$$\underline{p}(x,\gamma) = \frac{\partial f_c^e(x, \underline{\eta}_c(\gamma))}{\partial\, \underline{\eta}_c(\gamma)} = \begin{bmatrix} \frac{\partial f_c^e(x, \underline{\eta}_c(\gamma))}{\partial\, w(\gamma)} \\ \frac{\partial f_c^e(x, \underline{\eta}_c(\gamma))}{\partial\, m(\gamma)} \\ \frac{\partial f_c^e(x, \underline{\eta}_c(\gamma))}{\partial\, \sigma(\gamma)} \end{bmatrix}$$

with

$$\underline{p}(x,\gamma) = \begin{pmatrix} \frac{1}{w(\gamma)} \\ \frac{x-m(\gamma)}{(\sigma(\gamma))^2} \\ \frac{(x-m(\gamma))^2-(\sigma(\gamma))^2}{(\sigma(\gamma))^3} \end{pmatrix} f_c^e(x, \underline{\eta}_c(\gamma)) \ . \tag{6}$$

As a result, for this specific choice of continuous representation f_c^e, a closed-form expression for the matrix $\mathbf{P}(\gamma)$ can be obtained that is given by

$$\mathbf{P}(\gamma) = \begin{pmatrix} 1 & 0 & 0 \\ m(\gamma) & w(\gamma) & 0 \\ m^2(\gamma)+\sigma^2(\gamma) & 2\,w(\gamma)\,m(\gamma) & 2\,w(\gamma)\,\sigma(\gamma) \end{pmatrix} \ .$$

Solving for $\dot{\underline{\eta}}_c(\gamma)$ in (5) could be performed based on the matrix $\mathbf{P}(\gamma)$ directly. However, it is possible to calculate its inverse in closed form, which will be denoted by

$$\mathbf{Q}(\gamma) = \mathbf{P}(\gamma)^{-1} \ ,$$

with

$$\mathbf{Q}(\gamma) = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{m(\gamma)}{w(\gamma)} & \frac{1}{w(\gamma)} & 0 \\ \frac{m^2(\gamma)-\sigma^2(\gamma)}{2\,w(\gamma)\,\sigma(\gamma)} & -\frac{m(\gamma)}{w(\gamma)\,\sigma(\gamma)} & \frac{1}{2\,w(\gamma)\,\sigma(\gamma)} \end{pmatrix} \ .$$

Hence, we obtain

$$\dot{\underline{\eta}}_c(\gamma) = \mathbf{Q}(\gamma)\, \underline{b}(\gamma) \ , \tag{7}$$

with

$$\dot{\underline{\eta}}_c(\gamma) = [\dot{w}(\gamma),\, \dot{m}(\gamma),\, \dot{\sigma}(\gamma)]^T \ .$$

The right-hand-side of (7) that we will denote by $\underline{r}(\gamma)$ can now be written as

$$\begin{aligned} \underline{r}(\gamma) &= \mathbf{Q}(\gamma)\, \underline{b}(\gamma) \\ &= \mathbf{Q}(\gamma) \int_{\mathbb{R}} \underline{m} \cdot \tilde{f}^e(x,\gamma) \cdot \log\left(f_L(x)\right)\, dx \\ &= \int_{\mathbb{R}} \underline{q}(x,\gamma) \cdot \tilde{f}^e(x,\gamma) \cdot \log\left(f_L(x)\right)\, dx \end{aligned} \tag{8}$$

with

$$\underline{q}(x,\gamma) = \begin{pmatrix} 1 \\ \frac{x-m(\gamma)}{w(\gamma)} \\ \frac{(x-m(\gamma))^2-(\sigma(\gamma))^2}{2\,w(\gamma)\,\sigma(\gamma)} \end{pmatrix} \ .$$

Approximate Expressions for Vector $\underline{r}(\gamma)$: Calculating $\underline{r}(\gamma)$ in (8) amounts to calculating certain moments of the true posterior density $\tilde{f}^e(x,\gamma)$. As closed-form solutions are a rare occasion, we have to be content with approximate integration, which is pursued further here.

One viable option would be to replace the prior density $f^p(x)$ by its Dirac Mixture approximation and thus, turn the integration into summation. However, this is only efficient for small values of γ as long as the progressive Likelihood function is wide and does not modify the prior density too much. For larger γ, the Likelihood function typically becomes narrower and narrower and would force many Dirac components to zero, which is similar to the degeneration problem in particle filtering. By doing so, only very few Dirac components would really contribute to the integration, especially for larger γ, which is not desired.

Thanks to progressive processing, we have an approximate (Gaussian) posterior $f_c^e(x,\underline{\eta}_c(\gamma))$ available for every γ. This approximate posterior now allows to integrate only over those portions of the state space that contain the true posterior. For doing so, we rewrite (8) as

$$\underline{r}(\gamma) = \int_{\mathbb{R}} \underline{q}(x,\gamma) \cdot \underbrace{\frac{f_c^e(x,\underline{\eta}_c(\gamma))}{f_c^e(x,\underline{\eta}_c(\gamma))}}_{1} \cdot \tilde{f}^e(x,\gamma) \cdot \log\left(f_L(x)\right)\, dx$$

and replace the first occurrence of the continuous posterior $f_c^e(x,\underline{\eta}_c(\gamma))$ by its Dirac Mixture approximation f_d^e given by

$$f_d^e(x,\underline{\eta}_c(\gamma)) = \sum_{i=1}^{L_d} w_i(\underline{\eta}_c(\gamma))\ \delta(x - \hat{x}_i(\underline{\eta}_c(\gamma)))\ ,$$

with L_d the number of Dirac components. As a result, we obtain

$$\begin{aligned}\underline{r}(\gamma) &\approx \int_{\mathbb{R}} \underline{q}(x,\gamma) \cdot \frac{f_d^e(x,\underline{\eta}_c(\gamma))}{f_c^e(x,\underline{\eta}_c(\gamma))} \cdot \tilde{f}^e(x,\gamma) \cdot \log\left(f_L(x)\right)\, dx \\ &= \sum_{i=1}^{L_d} w_i \cdot \underline{q}(\hat{x}_i,\gamma)\, \frac{\tilde{f}^e(\hat{x}_i,\gamma)}{f_c^e(\hat{x}_i,\underline{\eta}_c(\gamma))} \cdot \log\left(f_L(\hat{x}_i)\right)\ ,\end{aligned} \tag{9}$$

where the explicit dependency of the parameters of the discrete density representation upon the parameters of the continuous part has been omitted.

It is important to note that the intermediate Gaussian posteriors $f_c^e(x,\underline{\eta}_c(\gamma))$ are used as proposal densities that should cover the support of the function to be integrated. By employing Dirac components solely in the important regions of the state space, all the components contribute to the integration. As a result, by far fewer components are required to achieve a good accuracy.

For the special case of a scalar measurement equation with additive Gaussian noise leading to a scalar Likelihood function

$$f_L(x) = \exp\left(-\frac{1}{2}\frac{(y-h(x))^2}{\sigma_v^2}\right)\ ,$$

we obtain

$$\log\left(f_L(\hat{x}_i)\right) = -\frac{1}{2}\frac{(y-h(\hat{x}_i))^2}{\sigma_v^2}\ .$$

for $i = 1, \ldots L_d$.

VI. Evaluation

For demonstrating the significant increase in performance achieved by the new filter, it is evaluated in comparison with the state-of-the-art filters. As a benchmark example, we consider the discrete-time cubic sensor problem

$$y = x^3 + v\ ,$$

where v is zero-mean Gaussian measurement noise. This is a canonical example, simple to understand, and, compared to some contrived artificial example, allows a good assessment of filter performance.

For comparison purposes, the proposed new filter is compared to

- the class of Linear Regression Kalman Filters (LRKF), specifically to
  - the Unscented Kalman Filter (UKF),
  - the Gaussian Filter (GF),
  - the Analytic LRKF,
- and an assumed-Gaussian density filter using moment matching for fitting a Gaussian to the true posterior employing Monte-Carlo integration for approximating the desired moments called Gaussian Particle Filter (GPF) [11].

For a measurement $\hat{y} = 3$, a prior Gaussian density with mean -1 and variance 1, and a cubic measurement equation corrupted by zero-mean Gaussian noise with variance 1.2, the results of one filter step for the different filters are shown in Figure 1. The true posterior of the Bayesian filter step is calculated on a fine grid with $30,000$ grid points. The ground truth for comparison purposes is the Gaussian density with mean and variance identical to the true posterior. A standard Unscented Kalman Filter (UKF) is employed. The Gaussian Filter (GF) is employed based on an optimal Dirac Mixture approximation of the prior Gaussian density according to [6], where the filter step is performed by assuming that measurement and state are jointly Gaussian. The Analytic LRKF computes the moments in closed form and additionally assumes that measurement and state are jointly Gaussian. The Gaussian Particle Filter (GPF) computes posterior mean and variance by means of 10^6 random samples drawn from the prior Gaussian density that are used for evaluating the product of Likelihood and prior density. Subsequently, the resulting weighted samples are used for moment matching. The new progressive Gaussian filter is applied with a density approximation using 30 Dirac components. To properly cover the support of the function to

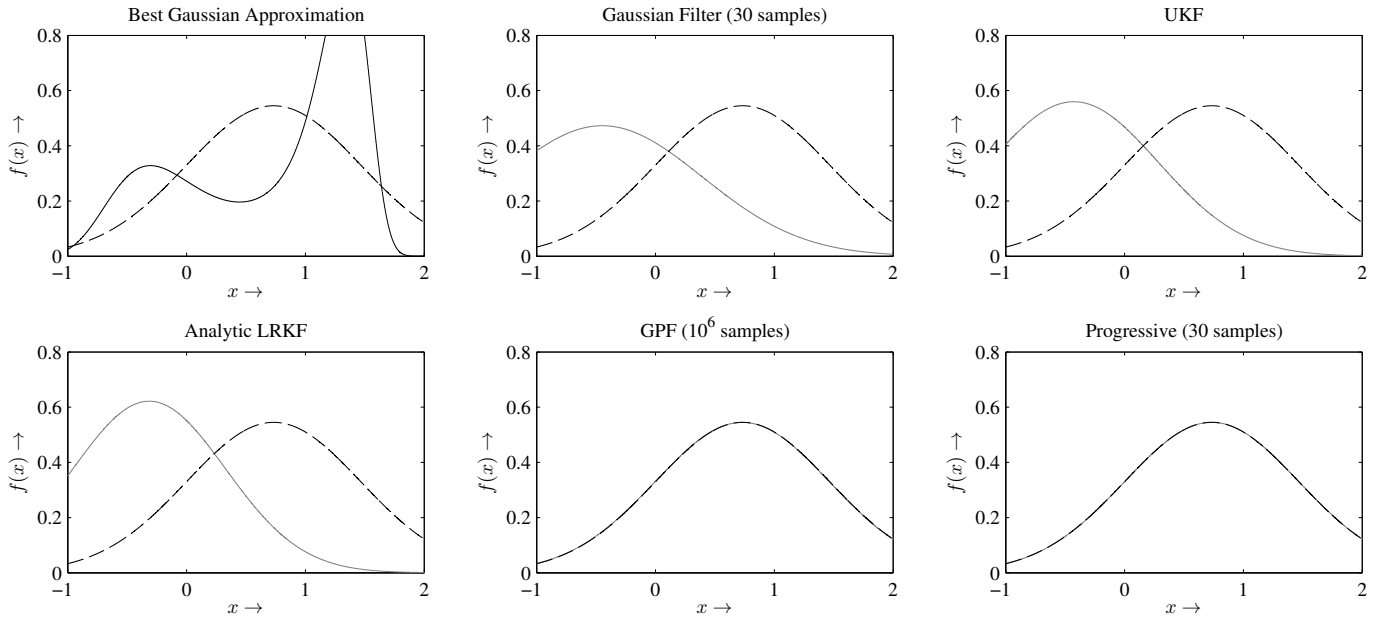

Figure 1. Comparison of the results for one filter step for the ground truth, i.e., the best Gaussian approximation computed numerically (always dashed lines), compared with the true posterior (solid black line), the Gaussian Filter (GF), the Unscented Kalman Filter (UKF), the Analytic LRKF, the Gaussian Particle Filter (GPF), and the new progressive Gaussian filter (all filters shown with solid gray lines).

be integrated in (8), the sample positions are scaled by a factor of 4.

Figure 1 clearly shows that the LRKFs have difficulties approximating the true posterior. In contrast, the GPF for 10^6 samples and the new progressive Gaussian filter produce results almost identical to the ground truth.

To illustrate the savings in the number of samples used by the new progressive Gaussian filter, its estimation quality is compared to the Gaussian Particle Filter (GPF) in Figure 2 for the same experimental parameters as above. It is obvious that very good estimation results are already obtained for very few (deterministic) samples of the Dirac Mixture approximation. In contrast, the results of the Gaussian particle filter are non-deterministic and the average convergence is much slower.

The results of recursive filtering over 50 times steps are shown in Figure 3. The recursion is started with a prior Gaussian density with mean -1 and variance 30. Noise mean is 0, the noise variance is 1.2. From time step 1 to 19, the true state is 1. At time step 20, the true state is changed to 0 and Gaussian noise with variance of 9 is added to all estimates.

The top plot in Figure 3 shows the estimated means of the true posterior, its best Gaussian approximation[6], the new progressive Gaussian filter, and the Analytic LRKF. The other variants of LRKF have been omitted as the Analytic LRKF yields the highest estimation quality.

The middle plot in Figure 3 shows the absolute mean error of the best Gaussian approximation, the new progressive filter, and the Analytic LRKF with respect to the mean of the true posterior.

The two top plots in Figure 3 show that the proposed progressive Gaussian filter provides results very close to the best Gaussian approximation, which itself is rather close to the mean of the true posterior. The LRKFs represented by their highest-quality variant, the Analytic LRKF, show a much larger deviation.

For solving the system of ordinary first-order differential equations, a standard solver (ODE15 in MATLAB) is employed. The number of steps taken for the inclusion of every measurement is shown in the bottom plot in Figure 3. Typically, only a few steps[7] are required. The number of steps is larger, when the state uncertainty is large, i.e., in the beginning and for the state change at time step 20, and so reflects the complexity of the estimation problem. It is important to note that the total number of function evaluations is given by the number of solver steps times the number of Dirac Mixture components in the discrete approximation of the intermediate posteriors. As the average number of solver steps is 43.4, the average number of function evaluations ($43.4 \cdot 30 = 1302$) is still very low.

[6]The best Gaussian approximation used as the ground truth at every time step is also recursively calculated on the grid based on the previous best Gaussian approximation. The accumulating error between the true posterior and the ground truth is unavoidable due to the Gaussian assumption.

[7]The median of the number of steps is 31.5.

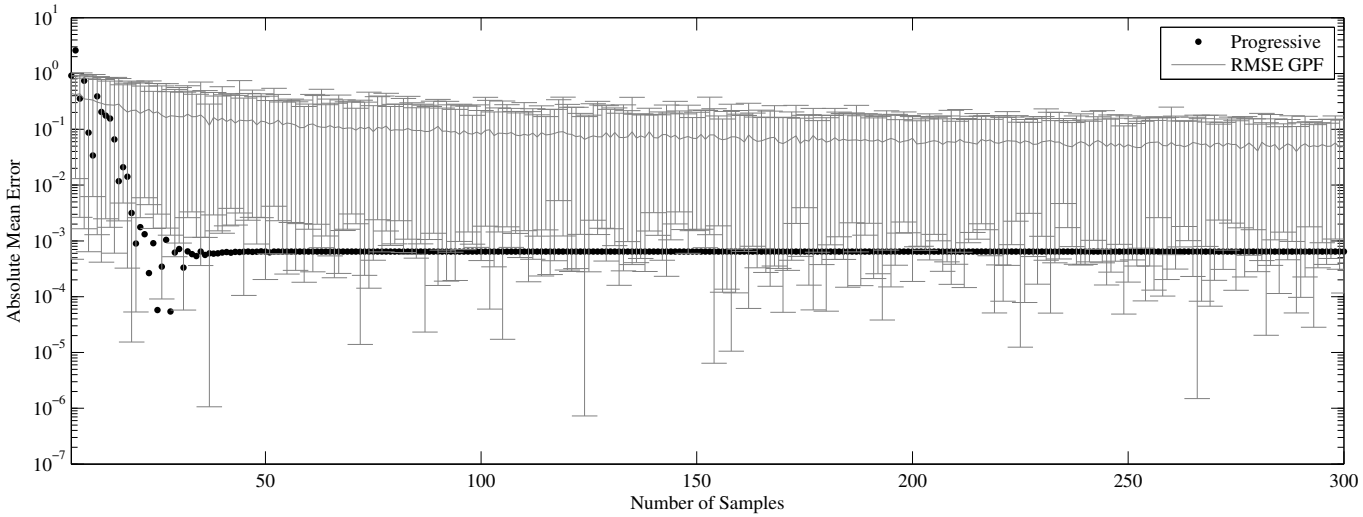

Figure 2. Comparison of the estimation quality as a function of the number of samples used between the Gaussian particle filter (GPF) and the proposed progressive Gaussian filter. As the GPF provides non-deterministic results, 100 Monte-Carlo runs have been performed to produce the RMSE and the min-max error bars.

VII. Conclusions

A progressive Gaussian filtering method has been introduced that calculates intermediate Gaussian posteriors and corresponding Dirac Mixture approximations. For the moment calculation, the Dirac Mixture approximations are employed to evaluate the true posterior density at discrete points only. In contrast to Monte-Carlo methods available for that purpose, the Dirac components all contribute to the integration as progressive processing is exploited to solely place components in the important regions of the state space. As a result, the new filter requires significantly less samples, is fast, efficient, and robust, and achieves a high estimation quality compared to state-of-the-art Gaussian filters.

For performing the filter step, the proposed method requires the Likelihood function and its logarithm. This is a disadvantage, as the Likelihood function is not always available. Current research work is investigating more advanced progressive processing schemes that do not explicitly require the Likelihood function.

References

[1] T. Lefebvre, H. Bruyninckx, and J. De Schutter, "The linear regression kalman filter," in *Nonlinear Kalman Filtering for Force-Controlled Robot Tasks*, ser. Springer Tracts in Advanced Robotics, 2005, vol. 19.

[2] ——, "Kalman filters for non-linear systems: a comparison of performance," *International Journal of Control*, vol. 77, no. 7, pp. 639–653, 2004.

[3] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

[4] S. J. Julier, "The scaled unscented transformation," in *American Control Conference, 2002. Proceedings of the 2002*, vol. 6. IEEE, 2002, pp. 4555– 4559 vol.6.

[5] D. Tenne and T. Singh, "The higher order unscented filter," in *American Control Conference, 2003. Proceedings of the 2003*, vol. 3. IEEE, Jun. 2003, pp. 2441– 2446 vol.3.

[6] 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)*, vol. 17, no. 2, Seoul, Republic of Korea, Jul. 2008.

[7] M. F. Huber, F. Beutler, and U. D. Hanebeck, "Semi-Analytic Gaussian Assumed Density Filter," in *Proceedings of the 2011 American Control Conference (ACC 2011)*, San Francisco, California, USA, Jun. 2011.

[8] F. Beutler, M. F. Huber, and U. D. Hanebeck, "Semi-Analytic Stochastic Linearization for Range-Based Pose Tracking," in *Proceedings of the 2010 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2010)*, Salt Lake City, Utah, USA, Sep. 2010.

[9] ——, "Gaussian Filtering using State Decomposition Methods," in *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*, Seattle, Washington, USA, Jul. 2009.

[10] 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.

[11] J. H. Kotecha and P. M. Djuric, "Gaussian particle filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2592– 2601, Oct. 2003.

[12] U. D. Hanebeck, K. Briechle, and A. Rauh, "Progressive Bayes: A New Framework for Nonlinear State Estimation," in *Proceedings of SPIE, AeroSense Symposium*, vol. 5099, Orlando, Florida, USA, May 2003, pp. 256 – 267.

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

[14] J. Hagmar, M. Jirstrand, L. Svensson, and M. Morelande, "Optimal parameterization of posterior densities using homotopy," in *2011 Proceedings of the 14th International Conference on Information Fusion (FUSION)*. IEEE, Jul. 2011, pp. 1–8.

[15] A. Rauh and U. D. Hanebeck, "Calculating Moments of Exponential Densities Using Differential Algebraic Equations," *IEEE Signal Processing Letters*, vol. 10, no. 5, pp. 144–147, May 2003.

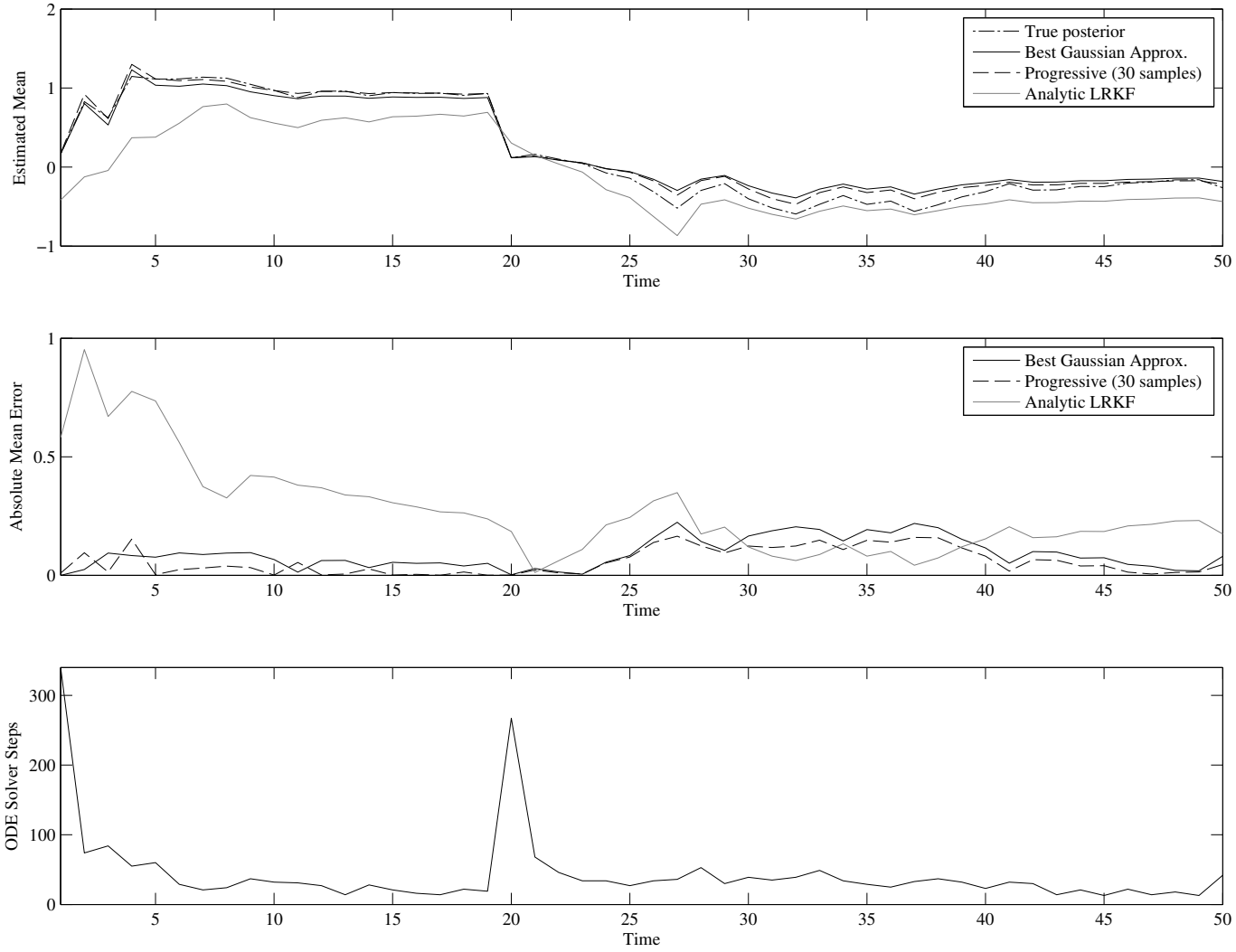

Figure 3. Results of recursive filtering over 50 times steps, where the recursion is started with a prior Gaussian density with mean -1 and variance 30. From time step 1 to 19, the true state is 1, from time step 20 to 50, the true state is changed to 0. Top: Estimated mean sequences. Middle: Absolute mean error of the best Gaussian approximation, the new progressive filter, and the Analytic LRKF with respect to the mean of the true posterior. Bottom: Number of steps taken for solving the ODE (7).

[16] M. F. Huber and U. D. Hanebeck, "Progressive Gaussian Mixture Reduction," in *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, Jul. 2008, pp. 1–8.

[17] F. Daum and J. Huang, "Nonlinear filters with particle flow induced by log-homotopy," *Proceedings of SPIE*, vol. 7336, no. 1, pp. 733 603–733 603–12, May 2009.

[18] U. D. Hanebeck, M. F. Huber, and V. Klumpp, "Dirac Mixture Approximation of Multivariate Gaussian Densities," in *Proceedings of the 2009 IEEE Conference on Decision and Control (CDC 2009)*, Shanghai, China, Dec. 2009.

[19] 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, USA, Dec. 2006.

[20] ——, "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, Sep. 2006, pp. 512–517.

[21] 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, USA, Dec. 2007, pp. 3065–3071.

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

[23] N. Oudjane and C. Musso, "Progressive correction for regularized particle filters," in *Information Fusion, 2000. FUSION 2000. Proceedings of the Third International Conference on*, vol. 2, Jul. 2000, pp. THB2/10 – THB2/17 vol.2.

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