Progressive Bayesian Estimation with Deterministic Particles

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

***Abstract*—This paper introduces an enhanced method for progressive Bayesian estimation based on a set of deterministic samples. The information of a given measurement is gradually introduced in order to avoid particle degeneration, which is usually encountered in standard particle filters. The main contribution of this paper is to derive a new method for exploiting smoothness assumptions about the unknown underlying density function of the state.**

I. Introduction

State estimation is the process of inferring the internal state of a system. Usually the state cannot be directly observed and must be inferred from available observations. When models for the time evolution of the state and for the relation between state and measurement are available, we have a model-based estimator. For discrete-time systems, a recursive estimator is usually employed that outputs a state estimate at every time step. The recursion includes two steps, a time update or prediction step for propagating the estimate through the system model and a measurement update or filter step for including the measurement information into the current estimate. Typically, both the measurements and the models suffer from uncertainties, so that an exact estimate of the state cannot be calculated. Rather, the estimate of the state must also include a measure of its uncertainty. In case of stochastic uncertainties, the uncertainty is represented by probability density functions. Hence, the stochastic filtering and prediction problem consists of the deterministic calculation of the posterior density function for the state given a prior density and densities describing the measurement and model uncertainties.

Exact solutions to the stochastic state estimation problem are rare and only exist for simple system such as the Kalman filter for linear systems with additive noise described by the first two moments [1]. For general nonlinear systems with complicated noise structures, either an approximation of the given system or an approximation of the occurring state density estimates has to be performed.

Methods for nonlinear state estimation can thus be classified according to the density type used for approximating the exact state density estimates. Typical choices are continuous densities such as the Gaussian density or Gaussian mixture densities. Another typical choice is the approximation of the exact state density estimate by a set of samples. This discrete representation is convenient as the samples can easily be propagated through nonlinear mappings and lead to posterior samples. This is in strong contrast to propagating a continuous density through a nonlinear system, which is usually complicated to do and almost always changes the type of the density. Sample representations are used in a Monte Carlo context by particle filters [2] or in a deterministic context such as quasi Monte Carlo methods, see [3].

In this paper, we consider the measurement update (filter step) and the time update (prediction step) for prior and posterior densities given by a set of samples. A measurement model and a system model is assumed to be known and a specific measurement is available at a certain time. It is further assumed that the measurement model and the given measurement have already been converted to a stochastic representation in the form of a likelihood function.

We consider deterministically placed (weighted) samples that we also call Dirac mixture density. These deterministic samples cover the interesting regions of the state space homogeneously, which is in strong contrast to random samples. Thus, far less samples are required. In addition, the samples are processed by deterministic methods, which guarantees repeatability: Repeating the same filter or prediction step gives exactly the same result. This is in contrast to standard particle filters [2], which employ randomized methods for performing the originally deterministic calculation of the posterior density functions.

We build upon the results in [4], a progressive Bayesian method for the measurement update. In order to avoid sample degeneration, the measurement step is not performed in one step, but gradually through several steps. Each step is carefully performed in such a way as to avoid degeneration. The approach was inspired by [5], which proposes a progressive measurement update in a Monte Carlo context. In contrast, [4] and this paper use a deterministic sample representation, which is deterministically moved by the update steps resulting in a flow of samples during the measurement update.

The paper is structured as follows. The next section, Sec. II, formulates the considered estimation problem with a focus on a single measurement update. The general method for attacking this estimation problem via progressive filtering is introduced in Sec. III. The involved challenges and the key idea are discussed in Sec. IV. Details on the implementation of the proposed new progressive estimator are given in Sec. V. Although the focus of this paper is on the measurement update, Sec. VI explains the corresponding prediction step, so that the proposed estimator can be used for recursively estimating the state of a dynamic system. In Sec. VII, the performance of the new progressive estimator is demonstrated in several simulations. Sec. VIII concludes the paper.

II. Problem Formulation

We are given a prior estimate of the state in the form of a Dirac mixture density $f^p(\underline{x})$ as

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

where w_i^p are the weights and $\underline{x}_i^p$ the locations of the L individual samples. For the prior Dirac mixture, we assume that the weights are all equal, i.e., $w_i^p = 1/L$ for $i = 1, \ldots, L$. In addition, we are given an arbitrary likelihood function $f^L(\underline{x})$.

We desire to calculate the resulting posterior density in the form of another Dirac mixture given by

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

Here, we also desire equal weights and we assume that the number of samples in the posterior equals the number of samples in the prior density. This is just for notational convenience, as of course the number of samples used for representing the uncertainty of the state can also be varied over time in order to reflect the time-varying complexity of its underlying density.

In theory, calculating the posterior density is done according to Bayes' law by multiplying the prior density with the likelihood function (and a subsequent appropriate renormalization). However, it is well known that in practical situations the naïve multiplication of a prior Dirac mixture with a narrow likelihood leads to so-called particle degeneration [6]. This means that all but a few particles are left with zero (or very small) weights so that the resulting particle density does not adequately represent the underlying density anymore.

Here, we will use the principle of progressive processing [7] in order to avoid particle degeneration as explained in the next section.

III. Progressive Filtering

The basic principle of the proposed new filter is as follows. The measurement update, i.e., the multiplication of the prior density with the likelihood function, is performed in several sub-steps, where each sub-step is carefully executed in order to avoid particle degeneration. For that purpose, the likelihood function $f^L(\underline{x})$ is decomposed into a set of S sub-likelihoods

$$f^L(\underline{x}) = \prod_{k=1}^{S} f_k^L(\underline{x}) \ . \tag{1}$$

There are many ways of calculating sub-likelihoods, as long as (1) holds. However, one convenient way is to use the following power decomposition

$$f^L(\underline{x}) = \left(f^L(\underline{x})\right)^{\Delta\gamma_1} \left(f^L(\underline{x})\right)^{\Delta\gamma_2} \cdots \left(f^L(\underline{x})\right)^{\Delta\gamma_S} \ , \tag{2}$$

with positive step sizes $\Delta\gamma_k > 0$ for $k = 1, \ldots, S$ and a total unit sum of step sizes $\Delta\gamma_1 + \Delta\gamma_2 + \cdots + \Delta\gamma_S = 1$.

The measurement update is then performed by calculating an intermediate posterior density by a reweighting of the samples representing the prior density according to

$$\tilde{f}_1^e(\underline{x}) = f^p(\underline{x}) \, f_1^L(\underline{x}) = f^p(\underline{x}) \left(f^L(\underline{x})\right)^{\Delta\gamma_1} \ .$$

This sub-posterior consisting of unequally weighted Dirac components is then replaced by an equally weighted Dirac mixture $f_1^e(\underline{x}) \approx \tilde{f}_1^e(\underline{x})$. This process is repeated by reweighting the corresponding sub-posterior again

$$\tilde{f}_k^e(\underline{x}) = f_{k-1}^e(\underline{x}) f_k^L(\underline{x}) = f_{k-1}^e(\underline{x}) \left(f^L(\underline{x})\right)^{\Delta\gamma_k} \tag{3}$$

and re-approximating them by unweighted Dirac mixtures

$$f_k^e(\underline{x}) \approx \tilde{f}_k^e(\underline{x}) \tag{4}$$

until all sub-likelihoods are used for $k = 2, \ldots, S$. The index k can be seen as an artificial time step. S artificial time steps are used for one complete measurement update. The final desired posterior is given by the last sub-posterior $f^e(\underline{x}) = f_S^e(\underline{x})$.

IV. Challenges and Key Idea

When naïvely performing the sub-update step in (3), we face the following challenge. Multiplying the current sub-posterior $f_{k-1}^e(\underline{x})$ with the current sub-likelihood $f_k^L(\underline{x})$ leads to varying weights between samples. This would be far more pronounced in case of one shot measurement updates and therefore leads to faster particle degeneration. However, even though we ensure a small variation between weights, there are still regions containing samples with large weights and regions containing samples with small weights.

When now subsequently performing the weight equalization in (4), this procedure would lead to a non-optimal lumping of the equalized samples in high-density regions. This can be intuitively grasped with a simple thought experiment: Consider a prior sample set with 10 samples and equal weights of 0.1. We assume a sub-likelihood that leads to two isolated samples with a weight of 0.15 each. The weight equalization procedure is now faced with the problem of replacing these two samples by three samples with weight 0.1 each. This is an ill-posed problem, with a non-unique result.

In [4], this problem is attacked by a regularization procedure, which essentially leads to a well-posed problem with a unique solution. Here, we proposed an easier and more efficient approach that assumes a smooth underlying density. It consists of smoothing the sub-posterior before performing the sub-update. Smoothing is performed by an adaptive up-sampling of the given sub-posterior, so we end up with an intermediate density comprising more samples. This ensures that a lot of narrow-spaced samples occur in high-density regions. After the sub-update, the up-sampled density is then re-approximated by a density with the original number of points.

V. Implementation

For implementing the progressive measurement update in (3) and (4), we use the following ingredients.

For up-sampling a given sub-posterior, we use a local method where individual samples are one-by-one replaced by several samples with smaller weights. This is suboptimal,

but computationally far less expensive than a global method that considers all original samples simultaneously. For doing so, we first estimate the value of the underlying continuous density at the sample point. This is described in Subsec. V-A.

Based on the value of the underlying density at the sample point, this sample point is replaced by a set of several precomputed deterministic samples. This procedure is described in Subsec. V-B.

Adequate step sizes for $\Delta\gamma_k$ have to be found at every artificial time step k that avoid too severe down-weighting of the sample points. Typically, these steps cannot be calculated beforehand, but have to be calculated online based on a given sub-posterior. Appropriate steps are derived in Subsec. V-C.

The new set of samples is now down-weighted by an appropriate sub-likelihood. So far, the sample positions have not been moved. Finally, the unequally weighted samples are replaced by a set of L equally weighted samples again, which leads to a motion of the samples towards regions of higher posterior density. This is described in Subsec. V-D.

A. Density Estimation

For replacing a single sample, we first estimate the value of the underlying density at this sample point. For density estimation, we use a k-nearest neighbor method. The number of nearest neighbors is set to one in this paper. Then, we directly replace the original sample by an appropriate sample set with R samples that are equally weighted. The weights of these samples are given by the weight of the original sample divided by R. To compensate for the smaller weights, the new samples are spaced more closely to each other than the original sample to its nearest neighbor.

For a given sub-posterior $f^e(\underline{x})$ represented by samples, we assume an underlying continuous density that we denote by $g^e(\underline{x})$, where we omit the index k of the sub-posterior. For estimating $g^e(\underline{x})$ at one of the sample points $\underline{x}_i^e$, its nearest neighbor $\underline{x}_j^e$ is calculated. The distance between the two points is denoted by d_{ij}. The density $g^e(\underline{x})$ at the component location is then estimated as the probability mass w_i^e of the sample point $\underline{x}_i^e$ divided by the volume V_i of the N-dimensional hypersphere with radius d_{ij}: $g^e(\underline{x}_i^e) = w_i^e / V_i$. The volume of the hypersphere is given by $V_i = c_N \, d_{ij}^N$, where c_N is the volume of the N-dimensional unit hypersphere.

B. Sample Replacement

Given an estimate of the value of the underlying density $g^e(\underline{x}_i^e)$ at a given sample point $\underline{x}_i^e$, this original sample is now replaced by a pre-calculated sample set with R samples.

For a given number of replacement samples R, we select an equally weighted Dirac mixture approximation for a standard normal distribution. In [8], an optimal sampling for Gaussian distributions based on the Localized Cumulative Distribution (LCD) was presented. A simplified version of this Gaussian sampling technique, reduced to the special case of standard normal distributions, was used to build a novel sample-based Kalman Filter, the Smart Sampling Kalman Filter (S^{2}KF) [9]. In order to further improve the quality of the sampling, a point symmetric version of the S^{2}KF is presented in [10]. An open-source implementation of the S^{2}KF and its point symmetric Gaussian sampling is also available online [11].

The sampling scheme for standard normal distributions is used for generating the replacement samples. This is done off-line and the samples are stored in a sample library. The resulting Dirac mixture density is called $f^R(\underline{x})$. Examples of $f^R(\underline{x})$ for different R are given in Fig. 1. For up-sampling, done on-line, the stored replacement Dirac mixture $f^R(\underline{x})$ is shifted to the position of the considered original sampling point $\underline{x}_i^e$ and scaled appropriately. For scaling, we first estimate the underlying continuous densities at every replacement sample of $f^R(\underline{x})$ with the method described in Subsec. V-A. Then, we scale $f^R(\underline{x})$ in such a way that the average over all R density values is equal to the underlying density $g^e(\underline{x}_i^e)$ at the considered sample point $\underline{x}_i^e$.

The up-sampled sub-posterior resulting from replacing all samples $\underline{x}_i^e$, $i = 1, \ldots, L$ of the sub-posterior $f_k^e(\underline{x})$ is denoted by $\hat{f}_k^e(\underline{x})$ and has $R \cdot L$ samples.

C. Using Sub-Likelihoods

As the number of sub-updates S and the specific values of $\Delta\gamma_k$ for $k = 1, \ldots, S$ in (2) typically cannot be calculated beforehand, we have to determine each step size $\Delta\gamma_k$ before performing a sub-update step. We consider the sub-update step at artificial time step k with a given sub-posterior $f_{k-1}^e(\underline{x})$ in the form of a Dirac mixture. The corresponding sub-likelihood $f_k^L(\underline{x}) = \left(f^L(\underline{x})\right)^{\Delta\gamma_k}$ is parametrized by $\Delta\gamma_k \in [0, 1]$. A small value of $\Delta\gamma_k$ means a "broad" sub-likelihood, while a large value of $\Delta\gamma_k$ leads to a "narrow" likelihood.

Our desire is now to find a value for $\Delta\gamma_k$ that leads to a certain minimum value for those specific Dirac mixture component that is down-weighted by the sub-likelihood the most. For finding this value of $\Delta\gamma_k$, we perform tentative sub-updates and normalize all weights in such a way, that the resulting maximum weight is equal to one (this is just done for determining the strength of the sub-update, finally the weights are normalized to sum to one). To minimize the number of evaluations, a bisection method is used. For the resulting smallest weight, one option is to a priori fix a certain value r, e.g., $r = 0.5$.

Another option in case of an adaptive replacement, i.e., a replacement with different numbers of replacement samples per original sample, is to use a down-weighting factor r that is inversely proportional to the maximum number of samples used for replacing a single sample. In that case the sample with the smallest weight would be associated with a replacement set of cardinality one, it stays the same, while the sample with the highest weight one would be associated with a replacement of a cardinality equal to the maximum number of replacement samples.

The $\Delta\gamma_k$ values are accumulated in a value γ for every artificial time step k as long as γ is less than one. When adding the calculated value of $\Delta\gamma_k$ to the current value of γ leads to a value of γ that is larger than one, $\Delta\gamma_k$ is set to $\Delta\gamma_k = 1 - \gamma$. This then is the last sub-update as now $\gamma = 1$ is reached.

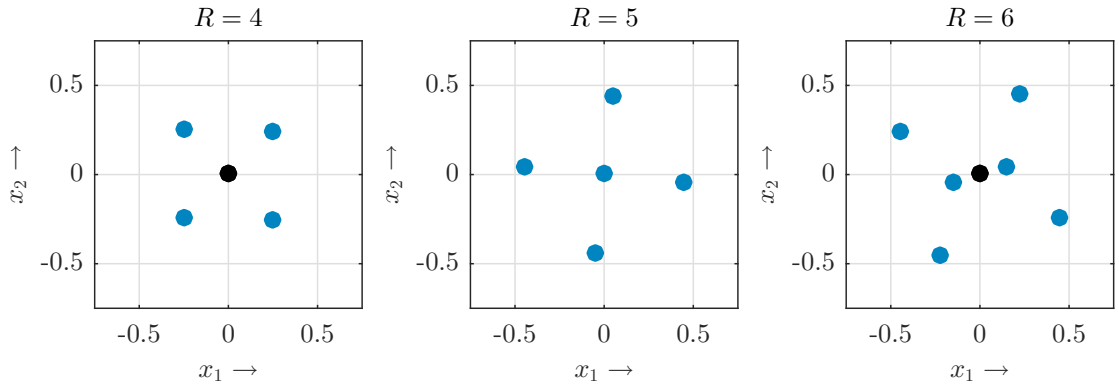

Fig. 1. Sample placement for $R = 4, 5$, and 6 samples of the up-sampling scheme (blue). Original sample placement shown in black.

D. From Nonequal Weights to Equal Weights

Given the slightly down-weighted density in (3) resulting from a sub-update step, the final step is to re-approximate this unequally weighted Dirac mixture with $R \cdot L$ components by an equally weighted one with L components. For this purpose, the optimal Dirac mixture reduction method from [12] is employed.

The complete Up-sampling Progressive Bayesian Filter (UsPBF) is summarized in Alg. 1.

VI. Prediction

In addition to the measurement equation, we now also consider a discrete-time, time-variant, nonlinear system equation for describing the time evolution of the system state given by

$$\underline{x}_{k'+1} = \underline{a}_{k'}(\underline{x}_{k'}, \underline{u}_{k'}, \underline{w}_{k'}) \tag{5}$$

with an arbitrary noise structure. The index k' here denotes real time steps of the system, not to be confused with the artificial time steps k used during a single progressive measurement update as detailed before.

$\underline{x}_{k'}$ is the system state st time step k'. $\underline{u}_{k'}$ is a known input, while $\underline{w}_{k'}$ is an unknown noise term acting upon the system. The density function describing the uncertainty of $\underline{w}_{k'}$ is assumed to be known and denoted by $f_{k'}^{w}(\underline{w}_{k'})$.

We now consider a set of samples $\underline{x}_{k',i}^{e}$, $i = 1, \ldots, L$ that result from the last measurement update, where we now included the time index k' that we omitted before as we so far focused on a single time step.

For performing the prediction, we draw M deterministic samples from the noise density $f_{k'}^{w}(\underline{w}_{k'})$ that we call $\underline{w}_{k',j}^{e}$, $j = 1, \ldots, M$. Now, the prediction step propagates all combinations of given state samples and noise samples through the nonlinear system equation (5) according to

$$\underline{x}_{k'+1,i,j}^{p} = \underline{a}_{k'}(\underline{x}_{k',i}^{e}, \underline{u}_{k'}, \underline{w}_{k',j}) \tag{6}$$

for $i = 1, \ldots, L$ and $j = 1, \ldots, M$, which yields $L \cdot M$ samples.

As the prediction step increases the number of samples, we re-approximate the resulting sample set with the Dirac mixture reduction method from [12]. This gives a sample set $\underline{x}_{k'+1,i}^{p}$, $i = 1, \ldots, L$ with cardinality L.

In the special case of additive noise, the prediction step is simplified. The system equation now reads

$$\underline{x}_{k'+1} = \underline{a}_{k'}(\underline{x}_{k'}, \underline{u}_{k'}) + \underline{w}_{k'} \ . \tag{7}$$

The total prediction is then done in two steps. First, we propagate the given samples $\underline{x}_{k',i}^{e}$, $i = 1, \ldots, L$ through the deterministic nonlinear part of the system equation, which gives L temporary predicted samples

$$\bar{\underline{x}}_{k'+1,i}^{p} = \underline{a}_{k'}(\underline{x}_{k',i}^{e}, \underline{u}_{k'})$$

for $i = 1, \ldots, L$.

In the second step, the temporarily predicted samples undergo a convolution with the noise samples as

$$\underline{x}_{k'+1,i,j}^{p} = \bar{\underline{x}}_{k'+1,i}^{p} + \underline{w}_{k',j}$$

for $i = 1, \ldots, L$ and $j = 1, \ldots, M$, which again yields $L \cdot M$ samples. In order to end up with the sample set $\underline{x}_{k'+1,i}^{p}$, $i = 1, \ldots, L$ with cardinality L, the Dirac mixture reduction method from [12] is used again.

When the proposed prediction step is used in conjunction with UsPBF, the complete estimator is called the Up-sampling Progressive Bayesian Estimator (UsPBE).

VII. Evaluation

We evaluate the proposed progressive Bayesian estimator including the up-sampling scheme in the context of linear and nonlinear two-dimensional estimation examples. First, we examine a basic example with linear system and measurement equation and compare the results to the Unscented Kalman Filter (UKF) [13] as a sanity check for the proposed filter. Second, we use a static setup to demonstrate the progression for the inclusion of a single measurement given a nonlinear measurement equation. Third, a dynamic system with a nonlinear measurement equation is considered that leads to multimodal state densities and clearly shows the advantages of the proposed new filter.

For all examples, we utilize 5 replacement samples for the up-sampled sub-posteriors, a number that gave satisfactory results in a wide variety of treated test cases.

Up-sampling Progressive Bayesian Filter (UsPBF)

Input : Prior sample set $\underline{x}_i^p$, $i = 1, \ldots, L$ representing the prior density $f^p(\underline{x})$, likelihood function $f^L(\underline{x})$, minimum down-weighting factor r, number of replacement samples R

Output : Posterior sample set $\underline{x}_i^e$, $i = 1, \ldots, L$ representing the posterior density $f^e(\underline{x})$

// Initialize time step
$k = 0$;
// Initialize sub-posterior for $k = 0$
$f_k^e(\underline{x}) = f^p(\underline{x})$;
while $\gamma < 1$ **do**
  // Increase time step
  $k = k + 1$;
  // Estimate underlying density at sample points
  d_{ij} is distance from sample point $\underline{x}_i^e$ to nearest neighbor $\underline{x}_j^e$;
  V_i is volume of hypersphere with radius d_{ij};
  Estimated density value given by $g^e(\underline{x}_i^e) = w_i^e / V_i$, $i = 1, \ldots, L$;
  // Replace original samples in sub-posterior $f^e(\underline{x})$
  Shift precomputed replacement density $f^R(\underline{x})$ to sample point position $\underline{x}_i^e$ and scale, $i = 1, \ldots, L$;
  Replace sample $\underline{x}_i^e$, $i = 1, \ldots, L$, resulting in $\hat{f}_{k-1}^e(\underline{x})$ with $R \cdot L$ equally weighted samples;
  // Calculate size of sub-update step
  Find $\Delta\gamma_k$ so that sub-likelihood $f_k^L(\underline{x})$ down-weights the samples no more than r;
  // Perform sub-update as in (3), but with up-sampled sub-posterior $\hat{f}_{k-1}(\underline{x})$
  $\tilde{f}_k^e(\underline{x}) = \hat{f}_{k-1}^e(\underline{x}) f_k^L(\underline{x})$, $\tilde{f}_k^e(\underline{x})$ has $R \cdot L$ unequally weighted samples;
  // Weight equalization
  Use the reduction method from [12] for $f_k^e(\underline{x}) \approx \tilde{f}_k^e(\underline{x})$ as in (4);
  // Limit γ to $[0, 1]$
  if $\gamma + \Delta\gamma > 1$ **then**
    $\Delta\gamma = 1 - \gamma$;
  end
end
// Extract desired final posterior density
$f^e(\underline{x}) = f_k^e(\underline{x})$ with $f_k^e(\underline{x})$ the last sub-posterior;

Algorithm 1: Up-sampling Progressive Bayesian Filter (UsPBF) for a single measurement update.

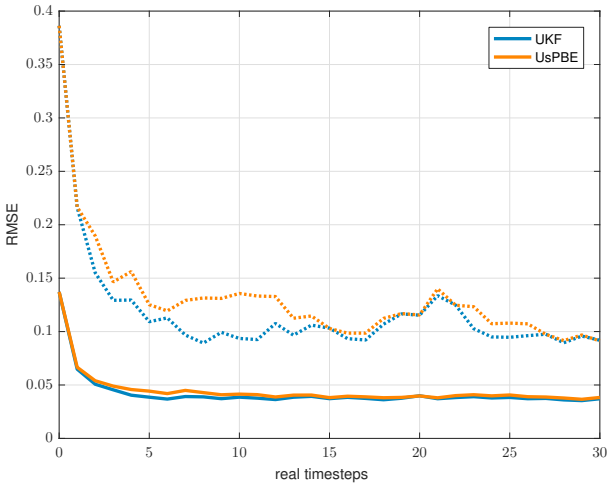

Fig. 2. Linear dynamic system. RMSE of the state mean and the maximal error bound (dotted) for the Unscented Kalman Filter (UKF) (blue) and the proposed Up-sampling Progressive Bayesian Estimator (UsPBE) (orange), utilizing 20 samples.

A. Linear Dynamic System

We estimate the system state $\underline{x}_{k'} = [x_{k',1}, x_{k',2}]^\top \in \mathbb{R}^2$, propagating according to a simple discrete-time linear dynamic system equation

$$\underline{x}_{k'} = \begin{bmatrix} 1.2 & 0 \\ 0 & 1.1 \end{bmatrix} \underline{x}_{k'-1} + \underline{w}_{k'}\,, \tag{8}$$

with additive zero-mean Gaussian system noise $\underline{w}_{k'} \sim \mathcal{N}([0,0]^\top, \mathrm{diag}(0.1^2, 0.1^2))$. The initial true system state is given by $\underline{x}_0 = [1,1]^\top$ with the initial estimate error covariance matrix $\mathbf{C}_0 = \mathrm{diag}(0.1^2, 0.1^2)$. 20 initial symmetric samples are drawn from a Gaussian $\mathcal{N}(\underline{x}_0, \mathbf{C}_0)$, generated according to the sampling scheme of the S^{2}KF [10].

The system state is observed through measurements $\underline{y}_{k'} \in \mathbb{R}^2$, the relation between system state and measurement given by the linear measurement equation

$$\underline{y}_{k'} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{x}_{k'} + \underline{v}_{k'}\,, \tag{9}$$

with additive zero-mean Gaussian measurement noise $\underline{v}_{k'} \sim \mathcal{N}([0,0]^\top, \mathrm{diag}(0.05^2, 0.05^2))$.

Fig. 2 shows the root mean square error (RMSE) for 30 real time steps, averaged over 100 runs. We can see estimation quality comparable to the UKF, the optimal filter for our test case. This clearly shows the advantages of the progressive measurement update, as this estimation quality is achieved with significantly less samples (in this case 20) than a particle filter with comparable quality would need.

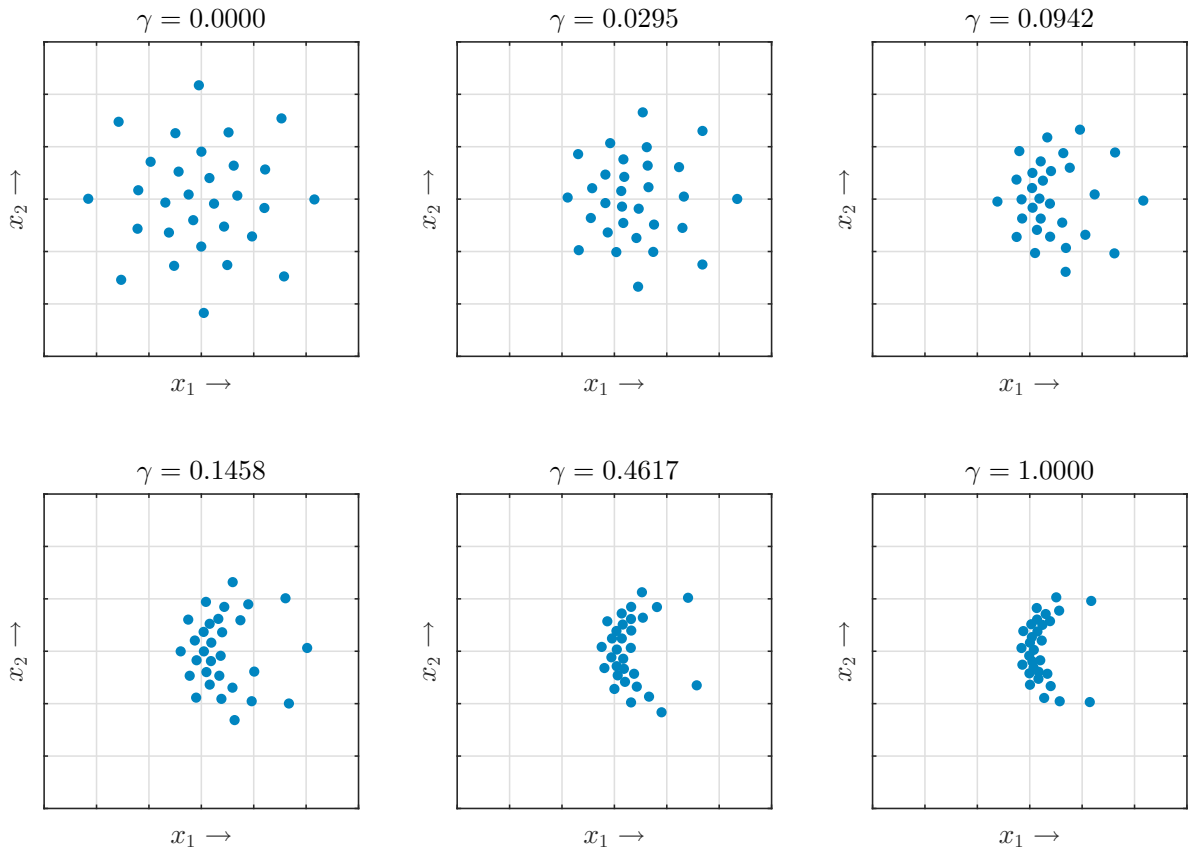

Fig. 3. Nonlinear static system. Single measurement update with the proposed Up-sampling Progressive Bayes Filter (UsPBF). The resulting sample placement for different values of γ is shown, which demonstrates the flow of particles.

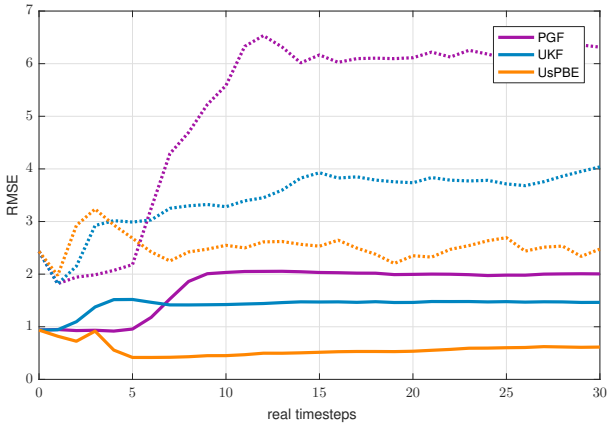

Fig. 4. Nonlinear dynamic system. RMSE of the state mean and the maximal error bound (dotted) for the Unscented Kalman Filter (UKF) (blue), Progressive Gaussian Filter (PGF) (purple), and the proposed Up-sampling Progressive Bayes Estimator (UsPBE) (orange), utilizing 40 samples.

B. Nonlinear Static System

We examine a nonlinear static system to illustrate the actual progression in the measurement update.

The initial system state is assumed to be $\underline{x}_0 = [0, 0]^\top$ and Gaussian distributed with covariance $\mathcal{N}([0, 0]^\top, \mathrm{diag}(1^2, 1^2))$. Again the S^{2}KF sampling scheme is used to draw 30 initial samples from this Gaussian. A scalar measurement equation

$$\begin{aligned} y_{k'} &= \left\| \underline{x}_{k'} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|_2 + v_{k'} \\ &= \sqrt{(x_{k',1} - 1)^2 + x_{k',2}^2} + v_{k'} \end{aligned} \tag{10}$$

is used, with additive zero-mean Gaussian noise $v_{k'} \sim \mathcal{N}(0, 0.1^2)$. The measurement received is $y_1 = 1$.

Fig. 3 shows the progression of the sample set for one measurement update for different values of γ. The samples progressively contract and approximate the shape of the posterior in an excellent fashion, allowing for the tracking of arbitrary densities.

C. Nonlinear Dynamic System

For the nonlinear test case, we model the system equation according to a vehicle moving with constant velocity in two dimensions, i.e.,

$$\underline{x}_{k'} = \underline{x}_{k'-1} + \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} + \underline{w}_{k'} \,. \tag{11}$$

Again the system noise is zero-mean Gaussian $\underline{w}_{k'} \sim \mathcal{N}([0, 0]^\top, \mathrm{diag}(0.1^2, 0.1^2)$. The initial estimate of the true system state $\underline{x}_0 = [-3, 0]^\top$ is assumed to be Gaussian distributed and small in the first and large in the second coordinate, giving a narrow initial estimate error covariance of the form $\mathrm{diag}(0.01^2, 0.9^2)$. This Gaussian is represented

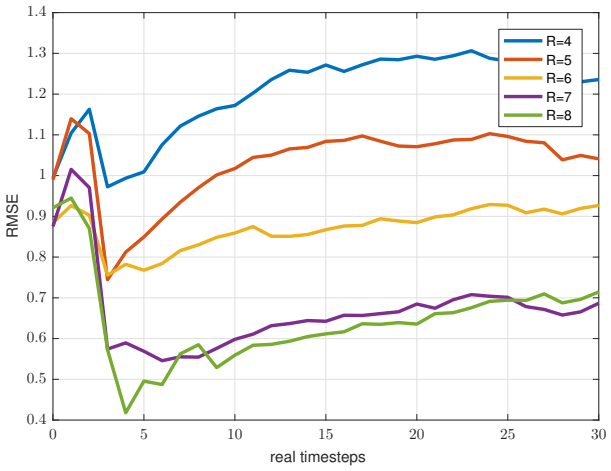

Fig. 5. Up-sampling. RMSE of the state mean for different numbers R of replacement samples used for up-sampling.

by 40 symmetric S^{2}KF samples. The constant velocity $[\Delta x_1, \Delta x_2]^\top$ is given by $[0.5, 0.1]^\top$.

The system state is measured through the scalar measurement equation

$$y_{k'} = \|\underline{x}_{k'}\|_2 + v_{k'} = \sqrt{x_{k',1}^2 + x_{k',2}^2} + v_{k'} \qquad (12)$$

with additive zero-mean Gaussian noise $v_{k'} \sim \mathcal{N}(0, 0.05^2)$. This amounts to measuring the vehicle position through the distance to a stationary sensor at the origin. As the measurement equation is a circular equation, there are two possible vehicle positions that result in the same measurement, making the problem bimodal.

Fig. 4 shows the RMSE for 30 real time steps, averaged over 100 runs, for the proposed progressive filter, the UKF, and the Progressive Gaussian Filter (PGF) [14]. The Upsampling Progressive Bayes Filter shows its strength in capturing both modes of the underlying model and following the true vehicle position. Both the UKF and PGF are Gaussian filters and thus have problems capturing the bimodal system. They often follow the "ghost" vehicle, leading to large estimation errors.

D. Resampling

Fig. 5 shows the influence of the replacement sampling on the estimation quality. Here, the RMSE for 30 real time steps, averaged over 100 runs is shown for different numbers of replacement samples (4-8) for the nonlinear dynamic system in VII-C.

We can see that higher samples numbers lead to better estimation results, as the state space coverage is improved. Higher numbers of replacement samples come with added cost of the subsequent weight equalization, thus, a balance between accuracy and computation time has to be struck. In our evaluation, 7 samples for the replacement are used, which gave satisfactory results.

VIII. Conclusions

This paper proposed several significant enhancements to the concept of progressive Bayesian estimation with deterministic sample sets as introduced in [4]. The major improvement is a new technique for exploiting smoothness assumptions about the underlying state densities by local up-sampling before progressive down-weighting is performed. All the necessary ingredients have been explained, including local density estimation, sample replacements, the calculation of appropriate sub-likelihoods, and the movement of samples by repetitiously replacing slightly down-weighted samples by equally weighted ones. As a result, a simple and robust mechanism for performing a nonlinear measurement update is achieved that gradually includes the given measurement information and lets the samples flow from a prior configuration to the desired posterior one.

References

[1] R. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME - Journal of Basic Engineering*, no. 82 (Series D), pp. 35–45, 1960.

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

[3] J. Spanier and E. Maize, "Quasi-Random Methods for Estimating Integrals Using Relatively Small Samples," *SIAM Review*, vol. 36, no. 1, pp. 18–44, Mar. 1994. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/1036002

[4] 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.

[5] 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.

[6] T. Li, S. Sun, T. P. Sattar, and J. M. Corchado, "Fighting Sample Degeneracy and Impoverishment in Particle Filters: A Review of Intelligent Approaches," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3944–3954, Jun. 2014, arXiv: 1308.2443. [Online]. Available: http://arxiv.org/abs/1308.2443

[7] 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.

[8] 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.

[9] J. Steinbring and U. D. Hanebeck, "LRKF Revisited: The Smart Sampling Kalman Filter (S^{2}KF)," *Journal of Advances in Information Fusion*, vol. 9, no. 2, pp. 106–123, Dec. 2014.

[10] J. Steinbring, M. Pander, and U. D. Hanebeck, "The Smart Sampling Kalman Filter with Symmetric Samples (accepted)," *Journal of Advances in Information Fusion*, 2016.

[11] J. Steinbring, "Nonlinear Estimation Toolbox." [Online]. Available: https://bitbucket.org/nonlinearestimation/toolbox

[12] U. D. Hanebeck, "Optimal Reduction of Multivariate Dirac Mixture Densities," *at - Automatisierungstechnik, Oldenbourg Verlag*, vol. 63, no. 4, pp. 265–278, Apr. 2015. [Online]. Available: http://dx.doi.org/10.1515/auto-2015-0005

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

[14] J. Steinbring and U. D. Hanebeck, "Progressive Gaussian Filtering Using Explicit Likelihoods," in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, Jul. 2014.