A Novel Gaussian Sum Filter Method for Accurate Solution to Nonlinear Filtering Problem

Gabriel Terejanu[a] Graduate Student terejanu@buffalo.edu

Puneet Singla[b] Assistant Professor psingla@buffalo.edu

Tarunraj Singh[b] Professor tsingh@buffalo.edu

Peter D. Scott[a] Associate Professor peter@buffalo.edu

[a]*Department of Computer Science & Engineering*

[b]*Department of Mechanical & Aerospace Engineering*

University at Buffalo, Buffalo, NY-14260.

***Abstract*—The paper presents two methods of updating the weights of a Gaussian mixture to account for the density propagation within a data assimilation setting. The evolution of the first two moments of the Gaussian components is given by the linearized model of the system. When observations are available, both the moments and the weights are updated to obtain a better approximation to the a posteriori probability density function. This can be done through a classical Gaussian Sum Filter. When the measurement model offers little or no information in updating the states of the system, better estimates may be obtained by updating the weights of the mixands to account for the propagation effect on the probability density function. The update of the forecast weights proves to be important in pure forecast settings, when the frequency of the measurements is low, when the uncertainty of the measurements is large or the measurement model is ambiguous making the system unobservable. Updating the weights not only provides us with better estimates but also with a more accurate probability density function. The numerical results show that updating the weights in the propagation step not only gives better estimates between the observations but also gives superior performance for systems where the measurements are ambiguous.**

Keywords: Gaussian Sum Filter, Fokker-Planck, Uncertainty Propagation, Weight Update, Improved Forecast.

I. Introduction

The nonlinear filtering problem has been extensively studied and various methods are provided in literature. The Extended Kalman Filter (EKF) is historically the first, and still the most widely adopted approach to solve the nonlinear estimation problem. It is based on the assumption that the nonlinear system dynamics can be accurately modeled by a first-order Taylor series expansion [1].

Since the EKF provides us only with a rough approximation of the a-posteriori probability density function (pdf) and solving for the exact solution of the conditional pdf is very expensive, researchers have been looking for mathematically convenient approximations. However, the accuracy and efficient implementation of these models have been an issue for highly nonlinear systems. The Bayes filter [2] offers the optimal recursive solution to the nonlinear filtering problem. However, the implementation of the Bayes filter in real time is computationally intractable because of the multi-dimensional integrals involved in the recursive equations. An efficient implementation of the Bayes filter is possible when all the involved random variables are assumed to be Gaussian [3], [4]. Several other approximate techniques such as Sequential Monte Carlo (SMC) methods [5], Gaussian closure [6] (or higher order moment closure), Equivalent Linearization [7], and Stochastic Averaging [8], [9] can also be used to find a solution to nonlinear filtering problem. Sequential Monte Carlo Methods or Particle filters [10] consists of discretizing the domain of the random variable into a set of finite number of particles and transforming these particle through the nonlinear map to obtain the distribution characteristics of the transformed random variable. Although SMC based methods are suitable for highly nonlinear and non-Gaussian models, they require extensive computational resources and effort, and becomes increasing infeasible for high-dimensional dynamic systems [11]. The latter methods (Gaussian closure, stochastic averaging etc.) are similar in several respects, and they are suitable only for linear or moderately nonlinear systems, because the effect of higher order terms can lead to significant errors. Furthermore, all these approaches provide only an approximate description of the uncertainty propagation problem by restricting the solution to a small number of parameters - for instance, the first N moments of the sought pdf.

In Ref. [12], a weighted sum of Gaussian density functions has been proposed to approximate the conditional pdf. It can be shown that as the number of Gaussian components increases the Gaussian sum approximation converges uniformly to any probability density function [13]. For a dynamical system with additive Gaussian white noise, the first two moments of the Gaussian components are propagated using the linearized model, and the weights of the new Gaussian components are set equal to the prior weights. In the case where observations are available both the moments and the weights are accordingly updated [12], [14] using Bayes rule to obtain an approximation of the a posteriori pdf, yielding the so called Gaussian Sum Filter (GSF). Extensive research has been done on Gaussian sum filters [10], which have become popular in the target tracking community. Such efforts have been materialized in methods like Gaussian Sum Filters [12], GSF with a more advanced measurement update [14], Mixture of Kalman Filters [15] and Interactive Multiple-Model [16]. However, in all of these methods *the weights of different components of a Gaussian mixture are kept constant while propagating the uncertainty through a nonlinear system and are updated only in the presence of measurement data.* This assumption is valid if the underlying dynamics is linear or the system is marginally nonlinear or measurements are precise and available very frequently. The same is not true for the general nonlinear case and new estimates of weights are required for accurate propagation of the state pdf. However, the existing literature provides no means for adaptation of the weights of different Gaussian components in the mixture model during the propagation of state pdf. The lack of adaptive algorithms for weights of Gaussian mixture are felt to be serious disadvantages of existing algorithms and provide the motivation for this paper.

The present paper is concerned with improving the conventional Gaussian sum filter approximation by adapting the weights of Gaussian mixture model in case of both pure propagation and measurement update. In the case of pure propagation, temporally sparse observations, large measurement noise or unobservable systems, a better approximation to the forecast pdf is obtained if the Gaussian components weights can be updated, in some optimal way, to account for the uncertainty propagation between the measurement time steps [17]. In Refs. [17], [18], two novel methods are discussed to update the weights corresponding to different components of the Gaussian mixture model. The first method updates the weights by constraining the Gaussian sum approximation to satisfy the Fokker-Planck-Kolmogorov (FPKE) equation and is appropriate for continuous-time dynamical systems. The second method updates the forecast weights such that they minimize the integral square difference between the true forecast pdf and its Gaussian sum approximation and is mostly appropriate for discrete-time nonlinear dynamical systems [17]. In this paper, we will make use of our earlier work on Gaussian sum approximation to improve upon conventional Gaussian sum filter methods and obtain a better approximation of a-posteriori pdf. We will present two new Gaussian sum filters for the nonlinear filtering problem and the performance of the proposed method will be compared with the performance of the classical Gaussian Sum Filter and SMC methods.

The organisation of the paper is as follows: first, the conventional Gaussian sum filter will be introduced in Section II followed by the introduction of two new methods to update the weights of Gaussian sum mixture during the propagation step. Numerical results are presented in Section V and the conclusions and future work are discussed in Section VI.

II. Conventional Gaussian Sum Filters

Consider a general n-dimensional continuous-time noise driven nonlinear dynamic system with uncertain initial conditions and discrete measurement model, given by the equations:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t)) + \mathbf{g}(t, \mathbf{x}(t))\Gamma(t) \quad (1)$$
$$\mathbf{z}_k = \mathbf{h}(t_k, \mathbf{x}_k) + \mathbf{v}_k \quad (2)$$

and a set of k observations, $\mathbf{Z}_k = \{\mathbf{z}_i \mid i = 1 \ldots k\}$.

We denote, $\mathbf{x}_k = \mathbf{x}(t_k)$, $\Gamma(t)$ represents a Gaussian white noise process with the correlation function $\mathbf{Q}\delta(t_{k+1} - t_k)$, and the initial state uncertainty is captured by the pdf $p(t_0, \mathbf{x}_0)$. The probability density function of the initial condition is given by the following Gaussian sum,

$$p(t_0, \mathbf{x}_0) = \sum_{i=1}^{N} w_0^i \mathcal{N}(\mathbf{x}_0 \mid \boldsymbol{\mu}_0^i, \mathbf{P}_0^i) \quad (3)$$

where,

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \mathbf{P}) = |2\pi\mathbf{P}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{P}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right]$$

The random vector $\mathbf{v}_k$ denotes the measurement noise, which is temporally uncorrelated, zero-mean random sequence with known covariance, $\mathbf{R}_k$. The process noise and the measurement noise are uncorrelated with each other and with the initial condition.

Let us assume that the underlying conditional pdf (forecast pdf if $t > t_k$ or a posteriori pdf if $t = t_k$) can be approximated by a finite sum of Gaussian pdfs

$$\hat{p}(t, \mathbf{x}(t) \mid \mathbf{Z}_k) = \sum_{i=1}^{N} w_{t|k}^i p_{g_i} \quad (4)$$
$$p_{g_i} = \mathcal{N}(\mathbf{x}(t) \mid \boldsymbol{\mu}_{t|k}^i, \mathbf{P}_{t|k}^i)$$

where $\boldsymbol{\mu}_{t|k}^i$ and $\mathbf{P}_{t|k}^i$ represent the conditional mean and covariance of the i^{th} component of the Gaussian pdf with respect to the k measurements, and $w_{t|k}^i$ denotes the amplitude of i^{th} Gaussian in the mixture. The positivity and normalization constraint on the mixture pdf, $\hat{p}(t, \mathbf{x}|\mathbf{Z}_k)$, leads to following constraints on the amplitude vector:

$$\sum_{i=1}^{N} w_{t|k}^i = 1, \quad w_{t|k}^i \geq 0, \quad \forall\ t \quad (5)$$

A Gaussian Sum Filter [12] may be used to propagate and update the conditional pdf. Since all the components of the mixture pdf (4) are Gaussian and thus, only estimates of their mean and covariance need to be propagated between t_k and t_{k+1} using the conventional Extended Kalman Filter time update equations:

$$\dot{\boldsymbol{\mu}}_{t|k}^i = \mathbf{f}(t, \boldsymbol{\mu}_{t|k}^i) \quad (6)$$
$$\dot{\mathbf{P}}_{t|k}^i = \mathbf{A}_{t|k}^i \mathbf{P}_{t|k}^i + \mathbf{P}_{t|k}^i (\mathbf{A}_{t|k}^i)^T + \mathbf{g}(t, \boldsymbol{\mu}_{t|k}^i)\mathbf{Q}\mathbf{g}^T(t, \boldsymbol{\mu}_{t|k}^i) \quad (7)$$
$$\mathbf{A}_{t|k}^i = \left.\frac{\partial \mathbf{f}(t, \mathbf{x}(t))}{\partial \mathbf{x}(t)}\right|_{\mathbf{x}(t)=\boldsymbol{\mu}_{t|k}^i} \quad (8)$$
$$w_{t|k}^i = w_{k|k}^i \quad \text{for} \quad t_k \leq t \leq t_{k+1} \quad (9)$$

The measurement update is done using Bayes rules:

$$p(\mathbf{x}_k|\mathbf{Z}_k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{k-1})}{\int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{k-1})\mathrm{d}\mathbf{x}_k} \tag{10}$$

While both the state and the covariance matrix are updated using the Extended Kalman Filter measurement update equations, the weights are updated using also Bayes rule but under the assumption that $\mathbf{P}^i_{k+1|k} \to 0$ as is shown in Ref. [3].

$$\boldsymbol{\mu}^i_{k+1|k+1} = \boldsymbol{\mu}^i_{k+1|k} + \mathbf{K}^i_k\left(\mathbf{z}_k - \mathbf{h}(t, \boldsymbol{\mu}^i_{k+1|k})\right) \tag{11}$$

$$\mathbf{P}^i_{k+1|k+1} = \left(\mathbf{I} - \mathbf{K}^i_k\mathbf{H}^i_k\right)\mathbf{P}^i_{k+1|k} \tag{12}$$

$$\mathbf{H}^i_k = \left.\frac{\partial \mathbf{h}(t, \mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k = \boldsymbol{\mu}^i_{k+1|k}} \tag{13}$$

$$\mathbf{K}^i_k = \mathbf{P}^i_{k+1|k}\mathbf{H}^i_k\left(\mathbf{H}^i_k\mathbf{P}^i_{k+1|k}(\mathbf{H}^i_k)^t + \mathbf{R}_k\right)^{-1} \tag{14}$$

$$w^i_{k+1|k+1} = \frac{w^i_{k+1|k}\beta^i_k}{\sum_{i=1}^N w^i_{k+1|k}\beta^i_k} \tag{15}$$

where

$$\beta^i_k = \mathcal{N}\left(\mathbf{z}_k - \mathbf{h}(t, \boldsymbol{\mu}^i_{k+1|k}),\ \mathbf{H}^i_k\mathbf{P}^i_{k+1|k}(\mathbf{H}^i_k)^t + \mathbf{R}_k\right) \tag{16}$$

An optimal state estimate and corresponding error covariance matrix can be obtained by making use of the following relations:

$$\boldsymbol{\mu}_{t|k} = \sum_{i=1}^N w^i_{t|k}\boldsymbol{\mu}^i_{t|k} \tag{17}$$

$$\mathbf{P}_{t|k} = \sum_{i=1}^N w^i_{t|k}\left[\mathbf{P}^i_{t|k} + (\boldsymbol{\mu}^i_{t|k} - \boldsymbol{\mu}_{t|k})(\boldsymbol{\mu}^i_{t|k} - \boldsymbol{\mu}_{t|k})^T\right] \tag{18}$$

Notice that it is assumed that weights $w^i_{t|k}$ do not change between measurement updates. This assumption is valid if underlying dynamics is linear or the system is marginally nonlinear. The same is not true for the general nonlinear case and new estimates of weights are required for accurate propagation of the state pdf. The reason that the weights are not changed is because it is assumed that the covariances are small enough [12] such that the linearizations become representative for the dynamics around the means. This is particularly a problem, if the uncertainty in measurement model is large and measurements are not available frequently.

In practice this assumption may be easily violated resulting in a poor approximation of the forecast pdf. Practically, the dynamic system may exhibit strong nonlinearities and the total number of Gaussian components, needed to represent the pdf, may be restricted due to computational requirements. The existing literature provides no means for adaption of the weights of different Gaussian components in the mixture model during the propagation of the state pdf. The lack of adaptive means for updating the weights of Gaussian mixture are felt to be serious disadvantages of existing algorithms and provide the motivation for this paper.

III. Update Forecast Weights - Continuous-Time Nonlinear Dynamical System

In this section, we summarize a recently developed method to update the weights of different components of the Gaussian mixture (4) during propagation. It can be shown that the true state probability density function, $p(t, \mathbf{x}(t)|\mathbf{Z}_k)$, satisfies the FPKE (19) between observations [2], which is a second order partial differential equation in $p(t, \mathbf{x}(t)|\mathbf{Z}_k)$.

$$\frac{\partial}{\partial t}p(t, \mathbf{x}|\mathbf{Z}_k) = \mathcal{L}_{\mathcal{FP}}p(t, \mathbf{x}|\mathbf{Z}_k) \tag{19}$$

The main idea is to modify the weights of the approximate conditional pdf given by Gaussian mixture (4), $\hat{p}(t, \mathbf{x}|\mathbf{Z}_k)$, such that it satisfies the FPKE (19).

The FPKE error can be used as a feedback to update the weights of different Gaussian components in the mixture pdf. In another words, we seek to minimize the FPKE error under the assumption (4), (6) and (7).

The substitution (4) in (19) leads to

$$e(t, \mathbf{x}) = \frac{\partial \hat{p}(t, \mathbf{x}|\mathbf{Z}_k)}{\partial t} - \mathcal{L}_{\mathcal{FP}}(\hat{p}(t, \mathbf{x}|\mathbf{Z}_k)) \tag{20}$$

$$\mathcal{L}_{\mathcal{FP}} = \left[-\sum_{i=1}^n \frac{\partial \mathrm{D}^{(1)}_i(t, \mathbf{x})}{\partial x_i} + \sum_{i=1}^n\sum_{j=1}^n \frac{\partial^2 \mathrm{D}^{(2)}_{ij}(t, \mathbf{x})}{\partial x_i \partial x_j}\right] \tag{21}$$

$$\mathrm{D}^{(1)}(t, \mathbf{x}) = \mathbf{f}(t, \mathbf{x}) + \frac{1}{2}\frac{\partial \mathbf{g}(t, \mathbf{x})}{\partial \mathbf{x}}\mathbf{Q}\mathbf{g}(t, \mathbf{x}) \tag{22}$$

$$\mathrm{D}^{(2)}(t, \mathbf{x}) = \frac{1}{2}\mathbf{g}(t, \mathbf{x})\mathbf{Q}\mathbf{g}^{\mathrm{T}}(t, \mathbf{x}) \tag{23}$$

where $\mathcal{L}_{\mathcal{FP}}(\cdot)$ is the so called Fokker-Planck operator, and,

$$\frac{\partial \hat{p}(t, \mathbf{x}|\mathbf{Z}_k)}{\partial t} = \sum_{i=1}^N w^i_{t|k}\left[\frac{\partial p_{g_i}}{\partial \boldsymbol{\mu}^i_{t|k}}^T \dot{\boldsymbol{\mu}}^i_{t|k} + \sum_{j=1}^n\sum_{k=1}^n \frac{\partial p_{g_i}}{\partial P^{i,jk}_{t|k}}\dot{P}^{i,jk}_{t|k}\right] \tag{24}$$

where $P^{i,jk}_{t|k}$ is the jk^{th} element of the i^{th} covariance matrix $\mathbf{P}^i_{t|k}$. Further, substitution of (22) and (23) along with (24) in (21) leads to

$$e(t, \mathbf{x}) = \sum_{i=1}^N w^i_{t|k}\mathfrak{L}_i(t, \mathbf{x}) = \mathfrak{L}^T\mathbf{w}_{t|k} \tag{25}$$

where $\mathbf{w}_{t|k}$ is a $N \times 1$ vector of Gaussian weights, and $\mathfrak{L}_i$ is given by

$$\begin{aligned}\mathfrak{L}_i(t, \mathbf{x}) = \Bigg[&\frac{\partial p_{g_i}}{\partial \boldsymbol{\mu}^i_{t|k}}^T \mathbf{f}(t, \boldsymbol{\mu}^i_{t|k}) + \sum_{j=1}^n\sum_{k=1}^n \frac{\partial p_{g_i}}{\partial P^{i,jk}_{t|k}}\dot{P}^{i,jk}_{t|k} \\ &+ \sum_{j=1}^n\Bigg(f_j(t, \mathbf{x})\frac{\partial p_{g_i}}{\partial x_j} + p_{g_i}\frac{\partial f_j(t, \mathbf{x})}{\partial x_j} \\ &+ \frac{1}{2}\frac{\partial d^{(1)}_j(t, \mathbf{x})p_{g_i}}{\partial x_j} - \frac{1}{2}\sum_{k=1}^n \frac{\partial^2 d^{(2)}_{jk}(t, \mathbf{x})p_{g_i}}{x_j x_k}\Bigg)\Bigg]\end{aligned} \tag{26}$$

$d^{(1)}(t, \mathbf{x})$ and $d^{(2)}(t, \mathbf{x})$ are given as:

$$d^{(1)}(t, \mathbf{x}) = \frac{1}{2}\frac{\partial \mathbf{g}(t, \mathbf{x})}{\partial \mathbf{x}}\mathbf{Q}\mathbf{g}(t, \mathbf{x}) \tag{27}$$

$$d^{(2)}(t, \mathbf{x}) = \frac{1}{2}\mathbf{g}(t, \mathbf{x})\mathbf{Q}\mathbf{g}^T(t, \mathbf{x}) \tag{28}$$

Further, different derivatives in the above equation can be computed using the following analytical formulas:

$$\frac{\partial p_{g_i}}{\partial \boldsymbol{\mu}^i_{t|k}} = (\mathbf{P}^i_{t|k})^{-1}\left(\mathbf{x}-\boldsymbol{\mu}^i_{t|k}\right)p_{g_i}$$
$$\frac{\partial p_{g_i}}{\partial \mathbf{P}^i_{t|k}} = \frac{p_{g_i}}{2}(\mathbf{P}^i_{t|k})^{-1}\left[\left(\mathbf{x}-\boldsymbol{\mu}^i_{t|k}\right)\left(\mathbf{x}-\boldsymbol{\mu}^i_{t|k}\right)^T(\mathbf{P}^i_{t|k})^{-1}-\mathbf{I}\right]$$
$$\frac{\partial p_{g_i}}{\partial \mathbf{x}} = -(\mathbf{P}^i_{t|k})^{-1}\left(\mathbf{x}-\boldsymbol{\mu}^i_{t|k}\right)p_{g_i}$$
$$\frac{\partial^2 p_{g_i}}{\partial \mathbf{x}\mathbf{x}^T} = -(\mathbf{P}^i_{t|k})^{-1}\left[\mathbf{I}+\left(\mathbf{x}-\boldsymbol{\mu}^i_{t|k}\right)\frac{\partial p_{g_i}}{\partial \mathbf{x}}^T\right]p_{g_i}$$

Now, at a given time instant, after propagating the mean, $\boldsymbol{\mu}^i_{t|k}$ and the covariance, $\mathbf{P}^i_{t|k}$, of individual Gaussian elements using (6) and (7), we seek to update weights by minimizing the FPE equation error over some volume of interest V:

$$\min_{w^i_{t|k}} \ \frac{1}{2}\int_V e^2(t,\mathbf{x})\mathrm{d}\mathbf{x} + \sum_{i=1}^N (w^i_{t|k}-w^i_{k|k})^2 \tag{29}$$
$$\text{s.t} \ \sum_{i=1}^N w^i_{t|k} = 1$$
$$w^i_{t|k} \geq 0, \quad i=1,\cdots,N$$

Here, the second term in the cost function is introduced to penalize large variations in the weights between two time steps. Since, FPKE error (25) is linear in Gaussian weights, $w^i_{t|k}$, hence, the aforementioned problem can be written as a quadratic programming problem as shown in Table I. Where $\mathbf{1}_{N\times 1}$ is a vector of ones, $\mathbf{0}_{N\times 1}$ is a vector of zeros and $\mathbf{L}$ is given by

$$\mathbf{L} = \int_V \mathfrak{L}(\mathbf{x})\mathfrak{L}^T(\mathbf{x})\mathrm{d}\mathbf{x}$$
$$= \begin{bmatrix} \int_V \mathfrak{L}_1\mathfrak{L}_1\mathrm{d}\mathbf{x} & \int_V \mathfrak{L}_2\mathfrak{L}_1\mathrm{d}\mathbf{x} & \cdots & \int_V \mathfrak{L}_N\mathfrak{L}_1\mathrm{d}\mathbf{x} \\ \int_V \mathfrak{L}_1\mathfrak{L}_2\mathrm{d}\mathbf{x} & \int_V \mathfrak{L}_2\mathfrak{L}_2\mathrm{d}\mathbf{x} & \cdots & \int_V \mathfrak{L}_N\mathfrak{L}_2\mathrm{d}\mathbf{x} \\ \vdots & & \ddots & \vdots \\ \int_V \mathfrak{L}_1\mathfrak{L}_N\mathrm{d}\mathbf{x} & \int_V \mathfrak{L}_2\mathfrak{L}_N\mathrm{d}\mathbf{x} & \cdots & \int_V \mathfrak{L}_N\mathfrak{L}_N\mathrm{d}\mathbf{x} \end{bmatrix} \tag{30}$$

Notice, to carry out this minimization, we need to evaluate integrals involving Gaussian pdfs over volume V which can be computed exactly for polynomial nonlinearity and in general can be approximated by the Gaussian quadrature method.

The minimization problem will substitute the Eq. (9) whenever an estimate has to be computed, even between measurements.

The summary of the Gaussian Sum Filter with forecast weight update for the continuous-time nonlinear dynamical systems and discrete measurement model is presented in Table I.

Table I
GSF2 - FORECAST WEIGHTS UPDATE METHOD I - CONTINUOUS-TIME DYNAMICAL SYSTEMS

Continuous-time nonlinear dynamics:
$\dot{\mathbf{x}}(t) = \mathbf{f}(t,\mathbf{x}(t)) + \mathbf{g}(t,\mathbf{x}(t))\Gamma(t)$
Discrete-time measurement model:
$\mathbf{z}_k = \mathbf{h}(t_k,\mathbf{x}_k) + \mathbf{v}_k$
Propagation:
$\dot{\boldsymbol{\mu}}^i_{t|k} = \mathbf{f}(t,\boldsymbol{\mu}^i_{t|k})$ $\dot{\mathbf{P}}^i_{t|k} = \mathbf{A}^i_{t|k}\mathbf{P}^i_{t|k} + \mathbf{P}^i_{t|k}(\mathbf{A}^i_{t|k})^T + \mathbf{g}(t,\boldsymbol{\mu}^i_{t|k})\mathbf{Q}\mathbf{g}^T(t,\boldsymbol{\mu}^i_{t|k})$ $\mathbf{A}^i_{t|k} = \left.\frac{\partial \mathbf{f}(t,\mathbf{x}(t))}{\partial \mathbf{x}(t)}\right\vert_{\mathbf{x}(t)=\boldsymbol{\mu}^i_{t|k}}$ $\mathbf{w}_{t|k} = \arg\min_{\mathbf{w}_{t|k}} \ \frac{1}{2}\mathbf{w}^T_{t|k}\mathbf{L}\mathbf{w}_{t|k} + (\mathbf{w}_{t|k}-\mathbf{w}_{k|k})^T(\mathbf{w}_{t|k}-\mathbf{w}_{k|k})$ subject to $\mathbf{1}^T_{N\times 1}\mathbf{w}_{t|k} = 1$ $\mathbf{w}_{t|k} \geq \mathbf{0}_{N\times 1}$
Measurement-Update:
$\boldsymbol{\mu}^i_{k+1|k+1} = \boldsymbol{\mu}^i_{k+1|k} + \mathbf{K}^i_k\left(\mathbf{z}_k - \mathbf{h}(t,\boldsymbol{\mu}^i_{k+1|k})\right)$ $\mathbf{P}^i_{k+1|k+1} = \left(\mathbf{I}-\mathbf{K}^i_k\mathbf{H}^i_k\right)\mathbf{P}^i_{k+1|k}$ $\mathbf{H}^i_k = \left.\frac{\partial \mathbf{h}(t,\mathbf{x}_k)}{\partial \mathbf{x}_k}\right\vert_{\mathbf{x}_k=\boldsymbol{\mu}^i_{k+1|k}}$ $\mathbf{K}^i_k = \mathbf{P}^i_{k+1|k}\mathbf{H}^i_k\left(\mathbf{H}^i_k\mathbf{P}^i_{k+1|k}(\mathbf{H}^i_k)^t + \mathbf{R}_k\right)^{-1}$ $w^i_{k+1|k+1} = \frac{w^i_{k+1|k}\beta^i_k}{\sum_{i=1}^N w^i_{k+1|k}\beta^i_k}$ $\beta^i_k = \mathcal{N}\left(\mathbf{z}_k - \mathbf{h}(t,\boldsymbol{\mu}^i_{k+1|k}),\ \mathbf{H}^i_k\mathbf{P}^i_{k+1|k}(\mathbf{H}^i_k)^t + \mathbf{R}_k\right)$

IV. UPDATE FORECAST WEIGHTS - DISCRETE-TIME NONLINEAR DYNAMICAL SYSTEM

For discrete-time case, consider the following nonlinear dynamic system with uncertain initial conditions given by the Gaussian mixture pdf (3) and discrete measurement model:

$$\mathbf{x}_{k+1} = \mathbf{f}(t_k,\mathbf{x}_k) + \boldsymbol{\eta}_k \tag{31}$$
$$\mathbf{z}_k = \mathbf{h}(t_k,\mathbf{x}_k) + \mathbf{v}_k \tag{32}$$

and a set of k observations, $\mathbf{Z}_k = \{\mathbf{z}_i \mid i = 1\ldots k\}$. Here $\boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0},\mathbf{Q}_k)$ represents the process noise. The same assumptions on process and measurement noise are made as in the continuous-time case.

The Gaussian Sum Filter used to propagate the conditional pdf, $\hat{p}(t_k,\mathbf{x}_k|\mathbf{Z}_k)$ is very similar with the filter used in continuous-time case, except the time-update step which is given by:

$$\boldsymbol{\mu}^i_{k+1|k} = \mathbf{f}(k,\boldsymbol{\mu}^i_{k|k}) \tag{33}$$
$$\mathbf{P}^i_{k+1|k} = \mathbf{A}^i_k\mathbf{P}^i_{k|k}(\mathbf{A}^i_k)^T + \mathbf{Q}_k \tag{34}$$
$$\text{where} \quad \mathbf{A}^i_k = \left.\frac{\partial \mathbf{f}(k,\mathbf{x}_k)}{\partial \mathbf{x}_k}\right\vert_{\mathbf{x}_k=\boldsymbol{\mu}^i_{k|k}} \tag{35}$$
$$w^i_{k+1|k} = w^i_{k|k} \tag{36}$$

Since the measurement model is also discrete, the same measurement update step as in the previous section, Eq.(11)-(15), holds also for discrete-time dynamical systems.

The same argument is made here regarding the forecast weights, Eq.(36), that are not changed during propagation. A better approximation of the forecast pdf and a more accurate estimate may be obtained by updating these weights.

The true conditional pdf, $p(t_{k+1}, \mathbf{x}_{k+1}|\mathbf{Z}_k)$ is given by the Chapman-Kolmogorov equation:

$$p(t_{k+1}, \mathbf{x}_{k+1}|\mathbf{Z}_k) = \int p(t_{k+1}, \mathbf{x}_{k+1}|t_k, \mathbf{x}_k) p(t_k, \mathbf{x}_k|\mathbf{Z}_k) \mathrm{d}\mathbf{x}_k \quad (37)$$

The new weights may be obtained in the least square sense to minimize the following integral square difference between the true conditional density function and its approximation:

$$\begin{aligned} \min_{w^i_{k+1|k}} \quad & \frac{1}{2} \int \left| p(t_{k+1}, \mathbf{x}_{k+1}|\mathbf{Z}_k) - \hat{p}(t_{k+1}, \mathbf{x}_{k+1}|\mathbf{Z}_k) \right|^2 \mathrm{d}\mathbf{x}_{k+1} \quad (38) \\ \text{s.t.} \quad & \sum_{i=1}^N w^i_{k+1|k} = 1 \\ & w^i_{k+1|k} \geq 0, \quad i = 1, \cdots, N \end{aligned}$$

Here, no *a priori* structure is used for the true conditional density function $p(t_{k+1}, \mathbf{x}_{k+1}|\mathbf{Z}_k)$, it is completely unknown.

The final formulation of the optimization (38) may be posed in the quadratic programming framework as shown in Table II and solved using readily available solvers.

This optimization problem will replace Eq.(36). Here $\mathbf{1}_{N\times 1}$ is a vector of ones and $\mathbf{0}_{N\times 1}$ is a vector of zeros, $\mathbf{w}_{k+1|k} = [w^1_{k+1|k}\ w^2_{k+1|k}\ \ldots\ w^N_{k+1|k}]^T$ is the vector of forecast weights, $\mathbf{w}_{k|k} = [w^1_{k|k}\ w^2_{k|k}\ \ldots\ w^N_{k|k}]^T$ is the prior weight vector, and $\mathbf{M}_{N\times N}$ is a symmetric matrix given by:

$$\mathbf{M}_{N\times N} = \int \mathfrak{M}(\mathbf{x}_{k+1}) \mathfrak{M}^T(\mathbf{x}_{k+1}) \mathrm{d}\mathbf{x}_{k+1} \quad (39)$$

where $\mathfrak{M}$ is a $N \times 1$ vector that contains all the propagated Gaussian components at time $k+1$:

$$\mathfrak{M}(\mathbf{x}_{k+1}) = \begin{bmatrix} \mathcal{N}(\mathbf{x}_{k+1} \mid \boldsymbol{\mu}^1_{k+1|k}, \mathbf{P}^1_{k+1|k}) \\ \mathcal{N}(\mathbf{x}_{k+1} \mid \boldsymbol{\mu}^2_{k+1|k}, \mathbf{P}^2_{k+1|k}) \\ \cdots \\ \mathcal{N}(\mathbf{x}_{k+1} \mid \boldsymbol{\mu}^N_{k+1|k}, \mathbf{P}^N_{k+1|k}) \end{bmatrix} \quad (40)$$

Thus, the components of matrix $\mathbf{M}$ are easily given by the product rule of two Gaussian density functions which yields another Gaussian density function. By integrating the product we are left only with the normalization constant:

$$\begin{aligned} m_{ij} = & \left| 2\pi (\mathbf{P}^i_{k+1|k} + \mathbf{P}^j_{k+1|k}) \right|^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{\mu}^i_{k+1|k} - \boldsymbol{\mu}^j_{k+1|k})^T \right. \\ & \left. \times (\mathbf{P}^i_{k+1|k} + \mathbf{P}^j_{k+1|k})^{-1} (\boldsymbol{\mu}^i_{k+1|k} - \boldsymbol{\mu}^j_{k+1|k}) \right] \quad (41) \\ m_{ii} = & \left| 4\pi \mathbf{P}^i_{k+1|k} \right|^{-1/2} \quad (42) \end{aligned}$$

Using Chapman-Kolmogorov equation (37), we are able to derive the following expressions for the elements of matrix $\mathbf{N}_{N\times N}$:

$$\begin{aligned} n_{ij} = & \int \mathcal{N}(\mathbf{f}(t_k, \mathbf{x}_k) \mid \boldsymbol{\mu}^i_{k+1|k}, \mathbf{P}^i_{k+1|k} + \mathbf{Q}_k) \\ & \times \mathcal{N}(\mathbf{x}_k \mid \boldsymbol{\mu}^j_{k|k}, \mathbf{P}^j_{k|k}) \ \mathrm{d}\mathbf{x}_k \quad (43) \\ = & \mathrm{E}_{\mathcal{N}(\mathbf{x}_k|\boldsymbol{\mu}^j_{k|k}, \mathbf{P}^j_{k|k})} \left[\mathcal{N}(\mathbf{f}(t_k, \mathbf{x}_k) | \boldsymbol{\mu}^i_{k+1|k}, \mathbf{P}^i_{k+1|k} + \mathbf{Q}_k) \right] \quad (44) \end{aligned}$$

Table II
GSF3 - Forecast Weights Update Method II - Discrete-Time Dynamical Systems

Discrete-time nonlinear dynamics:
$\mathbf{x}_{k+1} = \mathbf{f}(t_k, \mathbf{x}_k) + \boldsymbol{\eta}_k$
Discrete-time measurement model:
$\mathbf{z}_k = \mathbf{h}(t_k, \mathbf{x}_k) + \mathbf{v}_k$
Propagation:
$\boldsymbol{\mu}^i_{k+1\|k} = \mathbf{f}(k, \boldsymbol{\mu}^i_{k\|k})$ $\mathbf{P}^i_{k+1\|k} = \mathbf{A}^i_k \mathbf{P}^i_{k\|k} (\mathbf{A}^i_k)^T + \mathbf{Q}_k$ $\mathbf{A}^i_k = \left. \frac{\partial \mathbf{f}(k, \mathbf{x}_k)}{\partial \mathbf{x}_k} \right\|_{\mathbf{x}_k = \boldsymbol{\mu}^i_{k\|k}}$ $\mathbf{w}_{k+1\|k} = \arg \min_{\mathbf{w}_{k+1\|k}} \ \frac{1}{2} \mathbf{w}^T_{k+1\|k} \mathbf{M} \mathbf{w}_{k+1\|k} - \mathbf{w}^T_{k+1\|k} \mathbf{N} \mathbf{w}_{k\|k}$ subject to $\mathbf{1}^T_{N\times 1} \mathbf{w}_{k+1\|k} = 1$ $\mathbf{w}_{k+1\|k} \geq \mathbf{0}_{N\times 1}$
Measurement Update:
$\boldsymbol{\mu}^i_{k+1\|k+1} = \boldsymbol{\mu}^i_{k+1\|k} + \mathbf{K}^i_k \left(\mathbf{z}_k - \mathbf{h}(t, \boldsymbol{\mu}^i_{k+1\|k}) \right)$ $\mathbf{P}^i_{k+1\|k+1} = \left(\mathbf{I} - \mathbf{K}^i_k \mathbf{H}^i_k \right) \mathbf{P}^i_{k+1\|k}$ $\mathbf{H}^i_k = \left. \frac{\partial \mathbf{h}(t, \mathbf{x}_k)}{\partial \mathbf{x}_k} \right\|_{\mathbf{x}_k = \boldsymbol{\mu}^i_{k+1\|k}}$ $\mathbf{K}^i_k = \mathbf{P}^i_{k+1\|k} \mathbf{H}^i_k \left(\mathbf{H}^i_k \mathbf{P}^i_{k+1\|k} (\mathbf{H}^i_k)^t + \mathbf{R}_k \right)^{-1}$ $w^i_{k+1\|k+1} = \frac{w^i_{k+1\|k} \beta^i_k}{\sum_{i=1}^N w^i_{k+1\|k} \beta^i_k}$ $\beta^i_k = \mathcal{N}\left(\mathbf{z}_k - \mathbf{h}(t, \boldsymbol{\mu}^i_{k+1\|k}),\ \mathbf{H}^i_k \mathbf{P}^i_{k+1\|k} (\mathbf{H}^i_k)^t + \mathbf{R}_k \right)$

The expectations (44) may be computed using Gaussian Quadrature, Monte Carlo integration or Unscented Transformation [19]. Let the weighted sigma points $\{\mathcal{W}^l_k, \mathcal{X}^l_k\}$, where $l = 1 \ldots L$, sample the a posteriori normal distribution $\mathcal{N}\left(\mathbf{x}_k | \boldsymbol{\mu}^j_{k|k}, \mathbf{P}^j_{k|k}\right)$. The transformed sigma points $\{\mathcal{W}^l_k, \mathcal{X}^l_{k+1}\}$ are obtained by using the Unscented Transformation [19]. Thus we can numerically compute n_{ij} using the following relation:

$$n_{ij} = \sum_{l=1}^L \mathcal{W}^l_k \mathcal{N}(\mathcal{X}^l_{k+1} | \boldsymbol{\mu}^i_{k+1|k}, \mathbf{P}^i_{k+1|k}) \quad (45)$$

While in lower dimensions the unscented transformation is mostly equivalent with Gaussian quadrature, in higher dimensions the unscented transformation is computational more appealing in evaluating integrals since the number of points grows only linearly with the number of dimensions. However, this comes with a loss in accuracy [20], hence the need for a larger number of points [21] to capture additional information.

In the case of linear transformation, $\mathbf{f}(t_k, \cdot) = \mathbf{F}_k$, and by manipulating Eq.(43), $m_{ij} = n_{ij}$, yielding $\mathbf{M} = \mathbf{N}$, hence no change will occur to the weights. However, the numerical approximations (45) made in computing the expectations (44) may not give a perfect match between the two matrices, thus deteriorating the overall results in this case. Hence, such an update scheme may be recommended only for nonlinear systems.

The summary of the Gaussian Sum Filter with forecast weight update for the discrete-time nonlinear dynamical systems is presented in Table II.

V. Numerical Results

A. Example 1

To evaluate the performance of the forecast weight update scheme we consider the following continuous-time dynamical system with uncertain initial condition and discrete measurement model given by (48):

$$\dot{x} = \sin(x) + \Gamma(t) \quad \text{where} \quad Q = 1 \tag{46}$$

$$z_k = x_k^2 + v_k \quad \text{where} \quad R = 1 \tag{47}$$

$$x_0 \sim 0.1\mathcal{N}(-0.2, 1) + 0.9\mathcal{N}(0.2, 1) \tag{48}$$

The moments of the two Gaussian components are propagated for eight seconds using (6) and (7), using a sampling time of $\Delta t = 0.25$ sec and measurements available every 1 sec. Both methods have been applied for this example and every time step the estimate (17) has been computed and compared with the truth. The estimates of the two methods have been compared with the classical Gaussian Sum Filter using the root mean square error, RMSE - Eq.(49), averaged over 100 runs. Every run a different truth and afferent measurements have been generated.

$$\text{RMSE}(t) = \sqrt{\frac{1}{R}\sum_{j=1}^{R}(x_t^j - \mu_{t|k}^j)^2} \tag{49}$$

Where x_t^j is the true value of the state at time t for the j^{th} run and R is the total number of runs. In Fig. 2 we have computed the RMSE for all three methods. It is clear from this plot that we are able to better estimate the state with the incorporation of the weight update schemes.

Due to the multimodal nature of the a posteriori pdf, the RMSE is not an appropriate performance measure to compare different filters. A Boostrap Particle Filter with 10,000 particles, has also been implemented in order to have a proxy for the true a posteriori pdf. In Fig.1 we can see that by updating the weights of the Gaussian Sum during propagation we can better capture the true a posteriori pdf. In addition to this, log probability of the particles was computed according to the following relationship:

$$L = \sum_{j=1}^{M} \log \sum_{i=1}^{N} w_i \mathcal{N}(\mathbf{x}^j \mid \boldsymbol{\mu}_i, \mathbf{P}_i) \tag{50}$$

Here, M is the total number of particles used in the Boostrap Particle Filter. The higher value of L represents a better pdf approximation. Fig.3 shows the plot of log probability of the particles with and without updating the weights of the Gaussian mixture. As expected, updating the weights of the Gaussian mixture leads to higher log probability of the particles. Hence, we conclude that the adaptation of the weights during propagation leads to a more accurate a posteriori pdf approximation than without weight update. Due to the squared form of the measurement model, and bimodal nature of the forecast pdf, our measurements are not able to offer sufficient information to choose one mode of the conditional pdf, maintaining is bimodal nature. In such situation an accurate propagation makes the difference in providing better estimates and a more accurate conditional pdf.

While updating the forecast weights we obtain more accurate estimate of the state, it can be shown [17] that we also obtain a better conditional density function that agrees with the bimodal nature of the underlying true pdf, which can be used in applications that require more than just estimates as in Eq. (17).

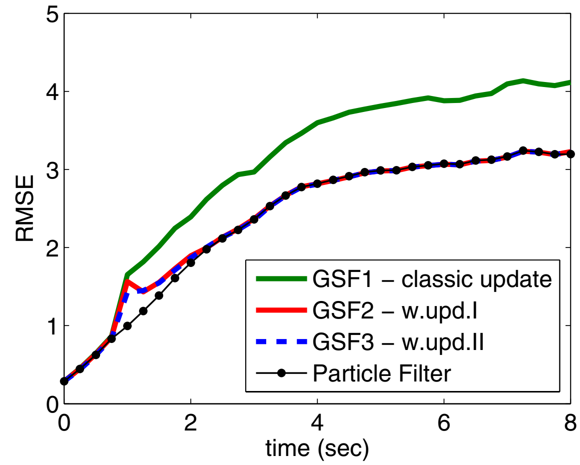


Figure 2. Example 1: RMSE Comparison (avg. over 100 runs)

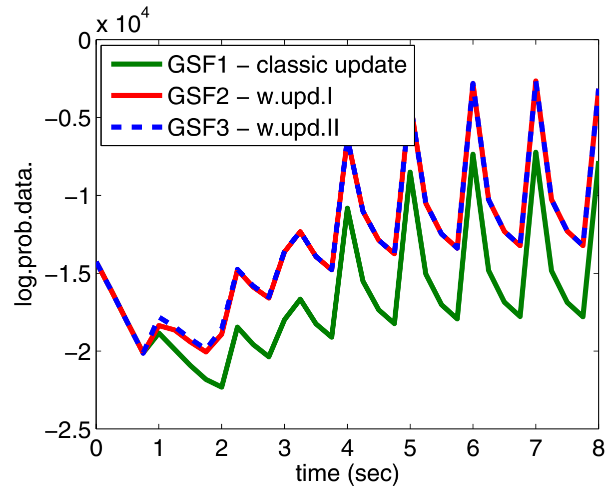
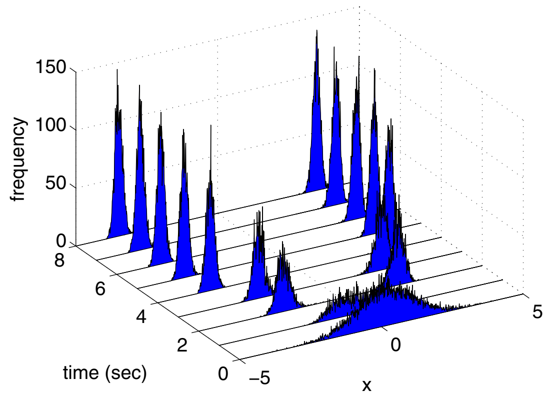
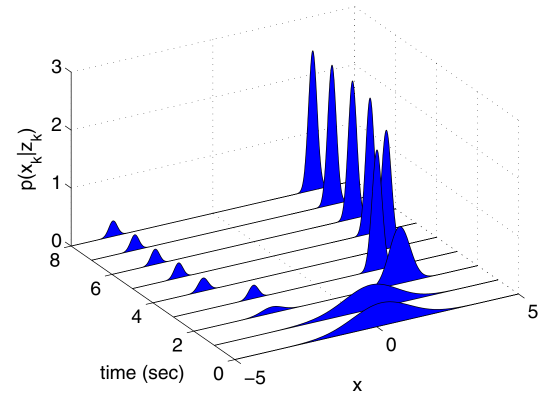


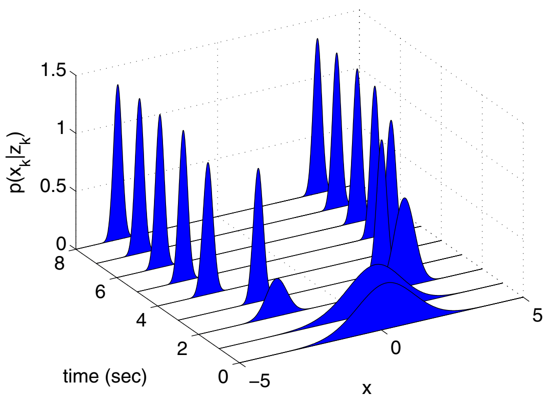
Figure 3. Example 1: Log probability of the particles (avg. over 100 runs)



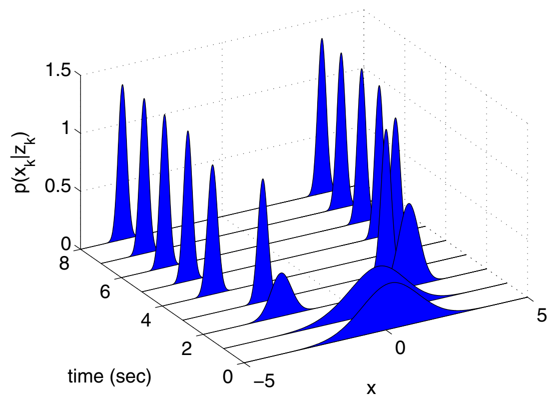
(a) Particle Filter - histogram a posteriori particles



(b) GSF1 classic update - a posteriori pdf



(c) GSF2 Table I - a posteriori pdf



(d) GSF3 Table II - a posteriori pdf

Figure 1. Example 1: A posteriori pdf comparison for one particular run

B. Example 2

We have also consider the following continuous-time dynamical system with uncertain initial condition and discrete measurement model given by (53):

$$\dot{x} = 5x(1 - x^2) + \Gamma(t) \quad \text{where} \quad Q = 0.25 \qquad (51)$$

$$z_k = x_k^2 + v_k \quad \text{where} \quad R = 0.01 \qquad (52)$$

$$x_0 \sim 0.1\mathcal{N}(-0.5, 2) + 0.9\mathcal{N}(0.5, 2) \qquad (53)$$

The moments of the two Gaussian components are propagated for 1 sec using (6) and (7), with a sampling time of $\Delta t = 0.01$ sec and measurements available every 0.1 sec. Both methods have been applied for this example and every time step the estimate (17) has been computed and compared with the truth.

In Fig.4 and Fig.5 we compare the three methods using the RMSE (49) and log probability of the particles (50) averaged over 100 runs. As in the previous example, a reduction in the

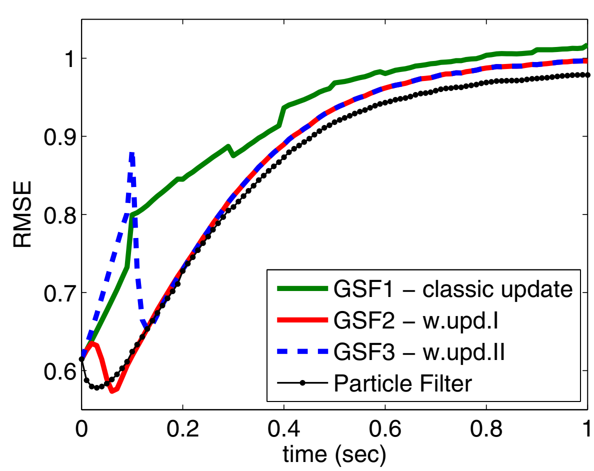


Figure 4. Example 2: RMSE Comparison (avg. over 100 runs)

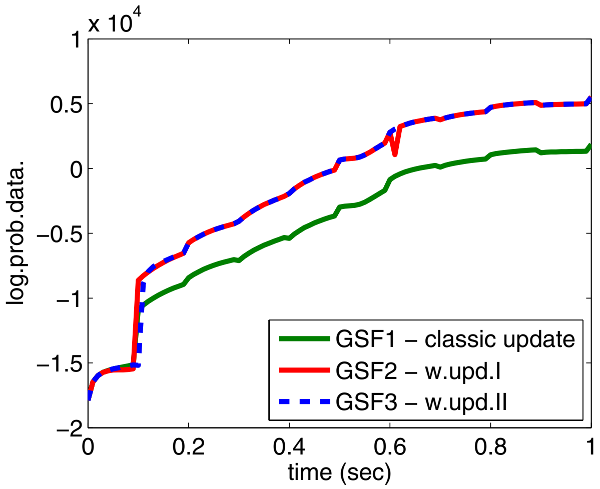


Figure 5. Example 2: Log probability of the particles (avg. over 100 runs)

error between the estimates and the truth is obtained and we better capture the *a posteriori* pdf when the forecast weights are updated during propagation.

VI. Conclusion

Two update schemes for the forecast weights are presented in order to obtain a better Gaussian sum approximation to the conditional pdf and a more accurate estimate. The difference between the two methods comes from the particularities of their derivations.

The first method updates the forecast weights such that the approximate conditional pdf, represented by a Gaussian sum, satisfies the Fokker-Planck equation. Since it deals with Fokker-Planck equation, the method is used with continuous-time nonlinear dynamical systems. The second method updates the weights such that they minimize the integral square difference between the true conditional probability density function and its approximation. The method is intended for discrete-time nonlinear dynamical systems but may be used also for continuous-time systems using their discrete representation. Both approaches are applied sequentially and the integrals used in the convex optimization have compact support and can be numerically approximated.

Two benchmark problems have been provided to compare the two methods with the usual procedure of not updating the weights in the propagation step. The results presented in this paper serve to illustrate the usefulness of adaptation of weights corresponding to different components of Gaussian sum model, providing compelling evidence and a basis for optimism. A detailed analysis of the convergence of the two methods is set as future work.

By updating the forecast weights, not only we obtain a more accurate estimate but also a better approximation to the conditional probability density function. The update methods presented are very useful when the measurement model offer limited or no information in updating the states of the system. Such situation arise when we are doing pure forecast, the frequency of the measurements is low, the uncertainty in the measurements is large or the measurement model is ambiguous, hence incapable of selecting one mode from the conditional probability density function. In these situations, when the measurements give little or no information in updating our uncertainty, a more accurate propagation of the conditional pdf is required.

Acknowledgment

This work was supported by the Defense Threat Reduction Agency (DTRA) under Contract No. W911NF-06-C-0162. The authors gratefully acknowledge the support and constructive suggestions of Dr. John Hannan of DTRA.

References

[1] J. Crassidis and J. Junkins, *Optimal Estimation of Dynamic Systems*. CRC Press, 2004.
[2] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, R. Bellman, Ed. Academic Press, 1970.
[3] B. D. Anderson and J. B. Moore, *Optimal Filtering*, T. Kailath, Ed. Prentice-Hall, 1979.
[4] A. Gelb, *Applied Optimal Estimation*. MIT Press, 1974.
[5] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte-Carlo Methods in Practice*. Springer-Verlag, April 2001.
[6] R. N. Iyengar and P. K. Dash, "Study of the random vibration of nonlinear systems by the gaussian closure technique," *Journal of Applied Mechanics*, vol. 45, pp. 393–399, 1978.
[7] J. B. Roberts and P. D. Spanos, *Random Vibration and Statistical Linearization*. Wiley, 1990.
[8] T. Lefebvre, H. Bruyninckx, and J. D. Schutter, "Kalman filters of non-linear systems: A comparison of performance," *International journal of Control*, vol. 77, no. 7, pp. 639–653, 2004.
[9] ——, "Comment on a new method for the nonlinear transformations of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, 2002.
[10] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter. Particle Filters for Tracking Applications*. Artech House, 2004.
[11] F. Daum and J. Huang, "Curse of dimensionality and particle filters," *Aerospace Conference, 2003. Proceedings. 2003 IEEE*, vol. 4, pp. 1979–1993, March 8-15, 2003.
[12] H. Alspach, D.; Sorenson, "Nonlinear bayesian estimation using gaussian sum approximations," *Automatic Control, IEEE Transactions on*, vol. 17, no. 4, pp. 439–448, 1972.
[13] H. W. Sorenson and D. L. Alspach, "Recursive bayesian estimation using gaussian sums," *Automatica*, vol. 7, pp. 465–479, 1971.
[14] K. Ito, K.; Xiong, "Gaussian filters for nonlinear filtering problems," *Automatic Control, IEEE Transactions on*, vol. 45, no. 5, pp. 910–927, 2000.
[15] R. Chen and J. Liu, "Mixture kalman filters," *Journal of the Royal Statistical Society: Series B*, vol. 62, pp. 493–508, 2000. [Online]. Available: citeseer.ist.psu.edu/chen00mixture.html
[16] Y. B. Shalom, X. R. Li, and T. Kirubarajan, *Estimatio with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., 2001.
[17] G. Terejanu, P. Singla, T. Singh, and P. D. Scott, "Uncertainty propagation for nonlinear dynamical systems using gaussian mixture models," *Journal of Guidance, Control, and Dynamics*, submitted for review Dec.2007.
[18] P. Singla and T. Singh, "A gaussian function network for uncertainty propagation through nonlinear dynamical system," *18th AAS/AIAA Spaceflight Mechanics Meeting, Galveston, TX*, Jan. 27-31, 2008.
[19] S. Julier and J. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
[20] A. Honkela, "Approximating nonlinear transformations of probability distributions for nonlinear independent component analysis," in *IEEE International Joint Conference on Neural Networks*, 2004.
[21] Y. Wu, M. Wu, D. Hu, and X. Hu, "An improvement to unscented transformation," in *17th Australian Joint Conference on Artificial Intelligence*, 2004.