Rao-Blackwellization for Adaptive Gaussian Sum Nonlinear Model Propagation

Sean R. Semper*

NASA Goddard Space Flight Center, Greenbelt, MD 20771

John L. Crassidis†

University at Buffalo, State University of New York, Amherst, NY 14260-4400

Jemin George‡

U.S. Army Research Laboratory, Adelphi, MD 20783

Siddharth Mukherjee§ and Puneet Singla¶

University at Buffalo, State University of New York, Amherst, NY 14260-4400

Introduction

When dealing with imperfect data and general models of dynamic systems, the best estimate is always sought in the presence of uncertainty or unknown parameters. In many cases, as the first attempt, the Extended Kalman filter (EKF) provides sufficient solutions to handling issues arising from nonlinear and non-Gaussian estimation problems. But these issues may lead unacceptable performance and even divergence. In order to accurately capture the nonlinearities of most real-world dynamic systems, advanced filtering methods have been created to reduce filter divergence while enhancing performance. Approaches, such as Gaussian sum filtering, grid based Bayesian methods and particle filters are well-known

*Aerospace Engineer, Guidance, Navigation and Control Hardware and Components Branch. Email: sean.r.semper@nasa.gov. Member AIAA.

†CUBRC Professor in Space Situational Awareness, Department of Mechanical & Aerospace Engineering. Email: johnc@buffalo.edu. Associate Fellow AIAA.

‡Engineer, Networked Sensing and Fusion Branch. Email: jemin.george.civ@mail.mil.

§Student, Department of Mechanical & Aerospace Engineering. Email: smukherj@buffalo.edu.

¶Associate Professor, Department of Mechanical & Aerospace Engineering. Email: psingla@buffalo.edu.

examples of advanced methods used to represent and recursively reproduce an approximation to the state probability density function (pdf). Some of these filtering methods were conceptually developed years before their widespread uses were realized. Advanced nonlinear filtering methods currently benefit from the computing advancements in computational speeds, memory, and parallel processing.

Grid based methods, multiple-model approaches and Gaussian sum filtering are numerical solutions that take advantage of different state coordinates or multiple-model methods that reduced the amount of approximations used. Choosing an efficient grid is very difficult for multi-dimensional state spaces, and oftentimes expensive computations must be done at each point. For the original Gaussian sum filter, a weighted sum of Gaussian density functions approximates the pdf but suffers at the update step for the individual component weight selections [1]. In order to improve upon the original Gaussian sum filter, Ref. [2] introduces a weight update approach at the filter propagation stage instead of the measurement update stage. This weight update is performed by minimizing the integral square difference between the true forecast pdf and its Gaussian sum approximation [2]. By adaptively updating each component weight during the nonlinear propagation stage an approximation of the true pdf can be successfully reconstructed.

Particle filtering (PF) methods have gained popularity recently for solving nonlinear estimation problems due to their straightforward approach and the processing capabilities mentioned above [3]. The basic concept behind PF is to represent any pdf as a set of random samples. As the number of samples increases, they will theoretically converge to the exact, equivalent representation of the desired pdf. When the estimated q^{th} moment is needed, the samples are used for its construction allowing further analysis of the pdf characteristics [4]. However, filter performance deteriorates as the dimension of the state vector increases. To overcome this problem Ref. [5] applies a marginalization technique for PF methods, decreasing complexity of the system to one linear and another nonlinear state estimation problem.

The marginalization theory was originally developed by Rao and Blackwell independently. According to Ref. [6] it improves any given estimator under every convex loss function. The improvement comes from calculating a conditional expected value, often involving integrating out a supportive statistic. In other words, Rao-Blackwellization allows for smaller but separate computations to be carried out while reaching the main objective of the estimator. In the case of improving an estimator's variance, any supporting statistic can be removed and its variance determined. Next, any other information that dependents on the supporting statistic is found along with its respective variance [6].

A new approach is developed here by utilizing the strengths of the adaptive Gaussian sum propagation in Ref. [2] and a marginalization approach used for PF methods found in

Ref. [7]. In the following sections a modified filtering approach is presented based on a special state-space model within nonlinear systems to reduce the dimensionality of the optimization problem in Ref. [2]. First, the adaptive Gaussian sum propagation is explained and then the new marginalized adaptive Gaussian sum propagation is derived. Finally, an example simulation is presented.

Gaussian Sum Nonlinear Model Propagation

In this section, the approach from Ref. [2] is used to approximate the pdf for nonlinear systems. The idea is to approximate the required posterior pdf by a *Gaussian mixture*, which is a weighted sum of Gaussian density functions [8]. According to Refs. [2,9], with a sufficient collection of Gaussian components, any pdf may be approximated as closely as desired. For this section the notation found in Ref. [2] is followed with minor differences in symbols, in order to present their new method. Given a pdf $p(\mathbf{x}_k)$, the Gaussian approximation is

$$p(\mathbf{x}_k) \approx \sum_{i=1}^{q} w_k^i \mathcal{N}(\mathbf{x}_k | \bar{\mathbf{x}}_k^i, P_k^i) \tag{1}$$

where the ith mean and covariance are denoted by $\bar{\mathbf{x}}_k^i$ and P_k^i, respectively. In order to have a valid pdf, the following requirement must be satisfied: that for some q and positive weights, w_i, $\sum_{i=1}^{q} w_k^i = 1$. Each Gaussian distribution, $\mathcal{N}(\mathbf{x}_k^i | \bar{\mathbf{x}}_k^i, P_k^i)$ is given by

$$\mathcal{N}(\mathbf{x}_k^i | \bar{\mathbf{x}}_k^i, P_k^i) = \frac{1}{\|2\pi P_k^i\|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_k^i)(P_k^i)^{-1}(\mathbf{x} - \bar{\mathbf{x}}_k^i)\right] \tag{2}$$

The pdf at time $k+1$ to be approximated as a Gaussian mixture is given by the Chapman-Kolmogorov equation [8]:

$$p(\mathbf{x}_{k+1}) = \int p(\mathbf{x}_{k+1} | \mathbf{x}_k) p(\mathbf{x}_k) \; d\mathbf{x}_k \tag{3}$$

where $p(\mathbf{x}_{k+1} | \mathbf{x}_k)$ is the probabilistic model of the state evolution (also known as *state transition pdf*). Consider the nonlinear model of the general form:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k) + \boldsymbol{\eta}_k \tag{4}$$

The state transition pdf depends on the pdf for the process noise variable $\boldsymbol{\eta}_k$, usually modeled as an additive zero-mean Gaussian noise processes with covariance Q_k [2]. Thus, the state

transition pdf can be written as

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) = \mathcal{N}\left(\mathbf{x}_{k+1}|\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k), Q_k\right) \tag{5}$$

In the traditional Gaussian sum approximation, the forecast density function, $p(\mathbf{x}_{k+1})$, is obtained by linearizing the nonlinear transformation and assuming that the weights of the different components are constant, i.e.,

$$\hat{p}(\mathbf{x}_{k+1}) = \sum_{i=1}^{q} w_{k+1}^i \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}_{k+1}^i, P_{k+1}^i) \tag{6}$$

where

$$w_{k+1}^i = w_k^i \tag{7a}$$

$$\bar{\mathbf{x}}_{k+1}^i = \mathbf{f}(\bar{\mathbf{x}}_{k+1}^i, \mathbf{u}_{k+1}^i, k+1) \tag{7b}$$

$$P_{k+1}^i = F_k(\bar{\mathbf{x}}^i) P_k^i F_k^T(\bar{\mathbf{x}}^i) + Q_k \tag{7c}$$

and

$$F_k(\bar{\mathbf{x}}^i) = \frac{\partial \mathbf{f}(\bar{\mathbf{x}}_k^i, \mathbf{u}_k^i, k)}{\partial \bar{\mathbf{x}}_k^i} \tag{8}$$

One of the advantages of the Gaussian sum approximation is, in a low-noise environment, the resulting estimator can be very nearly optimal. Conversely, the problem is, how to formulate an algorithmic procedure for on-line computation of updating the weights. This computation is difficult due to the fact that the number of components q can grow exponentially with time [2, 8]. Therefore, to obtain a favorable posterior pdf approximation developing better update laws for the future weights is necessary for enhancing performance.

Weight Update for Discrete-Time Dynamic Systems

While the traditional Gaussian sum approximation keeps the weights constant, Ref. [2] introduces a weight update scheme that utilizes the Chapman-Kolmogorov equation. That is, given the Gaussian sum approximation at some time k as

$$\hat{p}(\mathbf{x}_k) = \sum_{i=1}^{q} w_k^i \mathcal{N}(\mathbf{x}_k|\bar{\mathbf{x}}_k^i, P_k^i) \tag{9}$$

the Gaussian approximation as time $k+1$ may be written as

$$\hat{p}(\mathbf{x}_{k+1}) = \sum_{i=1}^{q} w_{k+1}^i \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}_{k+1}^i, P_{k+1}^i) \tag{10}$$

The approach in Ref. [2] finds the new weights of the Gaussian mixture such that the difference between the true pdf in Eq. (3) and the approximated pdf in Eq. (10) is minimized over the domain, $\mathbf{x}_{k+1}$, i.e.,

$$\min_{w^i_{k+1}} J = \frac{1}{2}\int \|p(\mathbf{x}_{k+1}) - \hat{p}(\mathbf{x}_{k+1})\|^2 \, d\mathbf{x}_{k+1} \tag{11a}$$

$$\text{s.t.} \quad \sum_{i=1}^{q} w^i_{i+1} = 1 \tag{11b}$$

$$w^i_{i+1} \geq 0, \quad i = 1, \ldots, q \tag{11c}$$

The terms in the above cost function can be expanded and rewritten as

$$J = \frac{1}{2} \int |p^2(\mathbf{x}_{k+1}) - 2\, p(\mathbf{x}_{k+1})\hat{p}(\mathbf{x}_{k+1}) + \hat{p}^2(\mathbf{x}_{k+1})| \, d\mathbf{x}_{k+1} \tag{12}$$

Substituting Eq. (10) and Eq. (3) gives

$$\frac{1}{2} \int \hat{p}^2(\mathbf{x}_{k+1}) \, d\mathbf{x}_{k+1} = \frac{1}{2} \int \left[\sum_{i=1}^{q} w^i_{k+1} \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1}) \right]^2 d\mathbf{x}_{k+1} \tag{13}$$

and

$$2p(\mathbf{x}_{k+1})\hat{p}(\mathbf{x}_{k+1}) = 2\int \left[\int p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_k) \, d\mathbf{x}_k \right] \left[\sum_{i=1}^{q} w^i_{k+1} \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1}) \right] d\mathbf{x}_{k+1} \tag{14a}$$

$$= 2\sum_{i=1}^{q} w^i_{k+1} \int \left[\int p(\mathbf{x}_{k+1}|\mathbf{x}_k) \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1}) d\mathbf{x}_{k+1} \right] p(\mathbf{x}_k) d\mathbf{x}_k \tag{14b}$$

which are quadratic and linear terms, respectively. Each linear term from the summation is denoted as y_i. Substituting Eq. (5) yields

$$y_i = \int \left[\int \mathcal{N}(\mathbf{x}_{k+1}|\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k), Q_k) \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1}) \, d\mathbf{x}_{k+1} \right] p(\mathbf{x}_k) \, d\mathbf{x}_k \tag{15a}$$

$$= \int \left[\mathcal{N}(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k)|\bar{\mathbf{x}}_{k+1}, P^i_{k+1} + Q_k) \right] p(\mathbf{x}_k) \, d\mathbf{x}_k \tag{15b}$$

The first term in the $p^2(\mathbf{x}_{k+1})$ expression is an additive constant and excluded from the optimization problem. Now the cost function can be written as

$$J = \frac{1}{2}\mathbf{w}^T_{k+1} M \mathbf{w}_{k+1} - \mathbf{w}^T_{k+1}\, \mathbf{y} \tag{16}$$

where $\mathbf{w}_{k+1} = [w^1_{k+1} \quad w^2_{k+1} \quad \cdots \quad w^q_{k+1}]^T$, $\mathbf{y}$ contains the components y_i, and $M \in \mathbb{R}^{q\times q}$ is a symmetric matrix given by

$$M = \int \mathfrak{M}(\mathbf{x}_{k+1})\mathfrak{M}^T(\mathbf{x}_{k+1})d\mathbf{x}_{k+1} \tag{17}$$

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

$$\mathfrak{M} = [\mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^1_{k+1}, P^1_{k+1}) \quad \mathcal{N}(\mathbf{x}_k|\bar{\mathbf{x}}^2_{k+1}, P^2_{k+1}) \quad \cdots \quad \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^q_{k+1}, P^q_{k+1})]^T \tag{18}$$

As a result the components of M are given by the product rule of two Gaussian density functions which yield another Gaussian density function [10]. Integrating the product leaves only the normalization constants performed. The off-diagonal elements of M are given by

$$\begin{aligned}
m_{ij} &= \int \mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1})\mathcal{N}(\mathbf{x}_{k+1}|\bar{\mathbf{x}}^j_{k+1}, P^j_{k+1})\ d\mathbf{x}_{k+1} \quad , i \neq j && \text{(19a)}\\
&= \mathcal{N}(\bar{\mathbf{x}}^i_{k+1}\big|\bar{\mathbf{x}}^j_{k+1}, P^i_{k+1} + P^j_{k+1})\\
&\times \int \mathcal{N}(\mathbf{x}^i_{k+1}\big|P^{ij}_{k+1}[(P^i_{k+1})^{-1}\bar{\mathbf{x}}^i_{k+1} + (P^j_{k+1})^{-1}\bar{\mathbf{x}}^j_{k+1}], P^{ij}_{k+1})\ d\mathbf{x}_{k+1}\\
&= \mathcal{N}(\bar{\mathbf{x}}^i_{k+1}|\bar{\mathbf{x}}^j_{k+1}, P^i_{k+1} + P^j_{k+1}) && \text{(19b)}\\
&= \|2\pi P^s_{k+1}\|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\bar{\mathbf{x}}^i_{k+1} - \bar{\mathbf{x}}^j_{k+1})^T(P^s_{k+1})^{-1}(\bar{\mathbf{x}}^i_{k+1} - \bar{\mathbf{x}}^j_{k+1})\right] && \text{(19c)}
\end{aligned}$$

where $P^s_{k+1} = P^i_{k+1} + P^j_{k+1}$ and $P^{ij} = [(P^i_{k+1})^{-1} + (P^j_{k+1})^{-1}]^{-1}$. The diagonal elements of M are given by

$$\begin{aligned}
m_{ii} &= \mathcal{N}(\bar{\mathbf{x}}^i_{k+1}|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1} + P^i_{k+1})\\
&= \|4\pi P^i_{k+1}\|^{-1/2}
\end{aligned} \tag{20}$$

Returning to the linear term $\mathbf{w}^T_{k+1}\,\mathbf{y}$, if the prior pdf is approximated by a Gaussian sum then

$$\begin{aligned}
y_i &= \int \left[\mathcal{N}(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k)|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1} + Q_k)\right] p(\mathbf{x}_k)\ d\mathbf{x}_k && \text{(21a)}\\
&\approx \int \left[\mathcal{N}(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k)|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1} + Q_k)\right] \hat{p}(\mathbf{x}_k)\ d\mathbf{x}_k && \text{(21b)}
\end{aligned}$$

Therefore,

$$y_i = \sum_{i=1}^{q} w^i_k \int \mathcal{N}(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k)|\bar{\mathbf{x}}^i_{k+1}, P^i_{k+1} + Q_k)\mathcal{N}(\mathbf{x}_k|\bar{\mathbf{x}}^j_k, P^j_k)\ d\mathbf{x}_k \tag{22a}$$

$$= \sum_{i=1}^{q} w_k^i N_{ij} \tag{22b}$$

The vector $\mathbf{w}_k = [w_k^1 \quad w_k^2 \quad \cdots \quad w_k^q]^T$ is the prior weight vector and the matrix $N \in \mathbb{R}^{q \times q}$ contains the following components:

$$N_{ij} = \int \mathcal{N}(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k) | \bar{\mathbf{x}}_{k+1}^i, P_{k+1}^i + Q_k) \mathcal{N}(\mathbf{x}_k | \bar{\mathbf{x}}_k^j, P_k^j) \, d\mathbf{x}_k \tag{23a}$$

$$= E_{\mathcal{N}(\mathbf{x}_k | \bar{\mathbf{x}}_k^j, P_k^j)} \left[\mathcal{N}(\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, k) | \bar{\mathbf{x}}_{k+1}^i, P_{k+1}^i + Q_k) \right] \tag{23b}$$

Note that an Unscented Transformation[a] (UT) can be used to compute the expectations of Eq. (23b). The final optimization problem is presented in the quadratic programming form as

$$\min J(w_{k+1}^i) = \frac{1}{2} \mathbf{w}_{k+1}^T M \mathbf{w}_{k+1} - \mathbf{w}_{k+1}^T N \mathbf{w}_k \tag{24a}$$

$$\text{s.t. } \mathbf{1}_{q \times 1}^T \mathbf{w}_{i+1}^i = 1 \tag{24b}$$

$$\mathbf{w}_{i+1}^i \geq \mathbf{0}_{q \times 1}, \quad i = 1, \ldots, q \tag{24c}$$

where $\mathbf{1}_{q \times 1}$ is a vector of ones and $\mathbf{0}_{q \times 1}$ is a vector of zeros.

Rao-Blackwellization Approach

A Gaussian sum filter approach allows an initial reduction of a computation complexity versus the classical Sequential Monte Carlo (SMC) method using the sequential importance sampling (SIS) algorithm. This is because a SMC method will approximate a continuous distribution of interest by a finite (but large) number of weighted random samples or *particles* in the state space. In theory any SIS method can approximate the posterior pdf of any form and solve any nonlinear system with any arbitrary distribution, i.e. the PF. Also, incorporating the familiar Bayesian approach known as Rao-Blackwellization, the SMC complexity can also be reduced by marginalizing out the conditional linear parts of the nonlinear model [6]. This results in a Rao-Blackwellized Particle Filter (RBPF) where the linear portion is estimated using a Kalman Fliter (KF), and the nonlinear portion using the original PF as mentioned earlier [7, 8, 11, 12]. In a similar goal of computational effort reduction for the RBPF, a Rao-Blackwellized Adaptive Gaussian Sum (RB-AGS) approach can allow a reduction of the number of Gaussian components by propagating linear portions using the KF and the nonlinear parts by the original EKF or Unscented Kalman Filter. Therefore, the

[a] See Ref. [2] for caveats concerning the UT.

Rao-Blackwellization and the new update laws will prove useful for some unique nonlinear systems.

Assume that the nonlinear system from Eq. (4) can be decomposed into a general state space model:

$$\mathbf{x}_{k+1}^{l} = \mathbf{f}_k^l(\mathbf{x}_k^n, \mathbf{u}_k) + \boldsymbol{\eta}_k^l \tag{25a}$$

$$\mathbf{x}_{k+1}^{n} = \mathbf{f}_k^n(\mathbf{x}_k^n, \mathbf{u}_k) + \boldsymbol{\eta}_k^n \tag{25b}$$

where the process noise $\boldsymbol{\eta}_k$ is

$$\boldsymbol{\eta}_k = \begin{bmatrix} \boldsymbol{\eta}_k^l \\ \boldsymbol{\eta}_k^n \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, Q_k), \quad Q_k = \begin{bmatrix} Q_k^l & 0 \\ 0 & Q_k^n \end{bmatrix} \tag{26}$$

Here, obtaining the pdf $p\left(\mathbf{x}_k\right) \equiv p\left(\mathbf{x}_k^l, \mathbf{x}_k^n\right)$ is desired. The pdf $p\left(\mathbf{x}_k\right)$ is calculated using the following two steps. First it is assumed, at time k, the Gaussian sum approximation of $p\left(\mathbf{x}_k^n\right)$ is given by

$$\hat{p}\left(\mathbf{x}_k^n\right) = \sum_{i=1}^{q} w_k^i \mathcal{N}\left(\mathbf{x}_k^n | \boldsymbol{\mu}_k^i, \Sigma_k^i\right) \tag{27}$$

for some chosen mean $\boldsymbol{\mu}_k^i$ and covariance Σ_k^i. Now the forecast approximation, $\hat{p}\left(\mathbf{x}_{k+1}^n\right)$, is calculated using the Chapman-Kolmogorov equation:

$$\hat{p}\left(\mathbf{x}_{k+1}^n\right) = \int p\left(\mathbf{x}_{k+1}^n | \mathbf{x}_k^n\right) \hat{p}\left(\mathbf{x}_k^n\right) d\mathbf{x}_k^n \tag{28}$$

Note that the precise expression for the conditional pdf, $p\left(\mathbf{x}_{k+1}^n | \mathbf{x}_k^n\right)$, can be written as

$$p\left(\mathbf{x}_{k+1}^n | \mathbf{x}_k^n\right) = \mathcal{N}\left(\mathbf{x}_{k+1}^n | \mathbf{f}_k^n(\mathbf{x}_k^n, \mathbf{u}_k), Q_k^n\right) \tag{29}$$

Following the same approach previously shown, the weights corresponding to individual Gaussian components can be obtained from solving the following optimization problem:

$$\min_{w_{k+1}^i} J = \frac{1}{2} \int \left\| p(\mathbf{x}_{k+1}^n) - \hat{p}(\mathbf{x}_{k+1}^n) \right\|^2 d\mathbf{x}_{k+1}^n \tag{30a}$$

$$\text{s.t.} \sum_{i=1}^{q} w_{i+1}^i = 1 \tag{30b}$$

$$w_{i+1}^i \geq 0, \quad i = 1, \ldots, q \tag{30c}$$

The above optimization problem can be written in vector format as

$$\min_{w^i_{k+1}} J(w^i_{k+1}) = \frac{1}{2}\mathbf{w}^T_{k+1} M \mathbf{w}_{k+1} - \mathbf{w}^T_{k+1} N \mathbf{w}_k \tag{31a}$$

$$\text{s.t. } \mathbf{1}^T_{q\times 1}\mathbf{w}^i_{i+1} = 1 \tag{31b}$$

$$\mathbf{w}^i_{i+1} \geq \mathbf{0}_{q\times 1}, \quad i = 1, \ldots, q \tag{31c}$$

where $\mathbf{w}_{k+1} = [w^1_{k+1} \quad w^2_{k+1} \quad \cdots \quad w^q_{k+1}]^T$, and $\mathbf{w}_k = [w^1_k \quad w^2_k \quad \cdots \quad w^q_k]^T$ is the prior weight vector. As before the symmetric Gaussian components for time $k+1$ are

$$M = \int \mathfrak{M}(\mathbf{x}^n_{k+1})\mathfrak{M}^T(\mathbf{x}^n_{k+1})\, d\mathbf{x}^n_{k+1} \tag{32}$$

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

$$\mathfrak{M} = [\mathcal{N}(\mathbf{x}^n_{k+1}|\boldsymbol{\mu}^1_{k+1}, \Sigma^1_{k+1}) \quad \mathcal{N}(\mathbf{x}^n_k|\boldsymbol{\mu}^2_{k+1}, \Sigma^2_{k+1}) \quad \cdots \quad \mathcal{N}(\mathbf{x}^n_{k+1}|\boldsymbol{\mu}^q_{k+1}, \Sigma^q_{k+1})]^T \tag{33}$$

Once again, for M the off-diagonal terms are

$$m_{ij} = \|2\pi\Sigma^s_{k+1}\|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\boldsymbol{\mu}^i_{k+1} - \boldsymbol{\mu}^j_{k+1})^T(\Sigma^s_{k+1})^{-1}(\boldsymbol{\mu}^i_{k+1} - \boldsymbol{\mu}^j_{k+1})\right] \tag{34}$$

where $\Sigma^s_{k+1} = \Sigma^i_{k+1} + \Sigma^j_{k+1}$ and the diagonal terms reduce to

$$\begin{aligned} m_{ii} &= \mathcal{N}(\boldsymbol{\mu}^i_{k+1}|\boldsymbol{\mu}^i_{k+1}, \Sigma^i_{k+1} + \Sigma^i_{k+1}) \\ &= \|4\pi\Sigma^i_{k+1}\|^{-1/2} \end{aligned} \tag{35}$$

Next the linear term $\mathbf{w}^T_{k+1}\mathbf{y}$ is made up of the individual components given by

$$y_i = \sum_{i=1}^{q} w^i_k \int \mathcal{N}(f^n_k(\mathbf{x}^n_k, \mathbf{u}_k, k)|\boldsymbol{\mu}_{k+1}, \Sigma^i_{k+1} + Q^n_k)\, \mathcal{N}(\mathbf{x}^n_k|\boldsymbol{\mu}^j_k, \Sigma^j_k)\, d\mathbf{x}^n_k \tag{36a}$$

$$= \sum_{i=1}^{q} w^i_k N_{ij} \tag{36b}$$

Now with $\mathbf{w}^T_{k+1}\,\mathbf{y} \equiv \mathbf{w}^T_{k+1} N \mathbf{w}_k$, each N_{ij} component can be seen as

$$N_{ij} = \int \mathcal{N}(f^n_k(\mathbf{x}^n_k, \mathbf{u}_k, k)|\boldsymbol{\mu}_{k+1}, \Sigma^i_{k+1} + Q^n_k)\, \mathcal{N}(\mathbf{x}^n_k|\boldsymbol{\mu}^j_k, \Sigma^j_k)\, d\mathbf{x}^n_k \tag{37a}$$

$$= E_{\mathcal{N}(\mathbf{x}^n_k|\boldsymbol{\mu}^j_k, \Sigma^j_k)}\left[\mathcal{N}(f^n_k(\mathbf{x}^n_k, \mathbf{u}_k, k)|\boldsymbol{\mu}_{k+1}, \Sigma^i_{k+1} + Q^n_k)\right] \tag{37b}$$

where the expectation above can be computed by the UT. Note that only the Gaussian

sum approximation is used for the marginalized nonlinear portion of special case state-space model.

The joint pdf $p\left(\mathbf{x}_k^l, \mathbf{x}_k^n\right)$ can be calculated as

$$p\left(\mathbf{x}_{k+1}^l, \mathbf{x}_{k+1}^n\right) = \int p\left(\mathbf{x}_{k+1}^l, \mathbf{x}_{k+1}^n | \mathbf{x}_k^n\right) p\left(\mathbf{x}_k^n\right) d\mathbf{x}_k^n \tag{38}$$

Note that since the conditional pdf $p\left(\mathbf{x}_{k+1}^l, \mathbf{x}_{k+1}^n | \mathbf{x}_k^n\right)$ is Gaussian and an analytical expression, it can be easily obtained as

$$p\left(\mathbf{x}_{k+1}^l, \mathbf{x}_{k+1}^n | \mathbf{x}_k^n\right) = \mathcal{N}\left(\begin{bmatrix} \mathbf{f}_k^l(\mathbf{x}_k^n, \mathbf{u}_k^n) \\ \mathbf{f}_k^n(\mathbf{x}_k^n, \mathbf{u}_k^n) \end{bmatrix}, \begin{bmatrix} Q_k^l & 0 \\ 0 & Q_k^n \end{bmatrix}\right) \tag{39}$$

Substituting the Gaussian sum approximation for $p\left(\mathbf{x}_k^n\right)$, an approximation for the joint pdf, $\hat{p}\left(\mathbf{x}_k\right)$, can be written as

$$\hat{p}\left(\mathbf{x}_{k+1}^l, \mathbf{x}_{k+1}^n\right) = \sum_{i=1}^{q} w_k^i \int \mathcal{N}\left(\begin{bmatrix} \mathbf{f}_k^l(\mathbf{x}_k^n, \mathbf{u}_k^n) \\ \mathbf{f}_k^n(\mathbf{x}_k^n, \mathbf{u}_k^n) \end{bmatrix}, \begin{bmatrix} Q_k^l & 0 \\ 0 & Q_k^n \end{bmatrix}\right) \mathcal{N}\left(\mathbf{x}_k^n | \boldsymbol{\mu}_k^i, \Sigma_k^i\right) d\mathbf{x}_k^n \tag{40}$$

The above integral can be considered as an expectation since

$$\int \mathcal{N}\left(\mathbf{f}_k(\mathbf{x}_k^n, \mathbf{u}_k^n), Q_k\right) \mathcal{N}\left(\mathbf{x}_k^n | \boldsymbol{\mu}_k^i, \Sigma_k^i\right) d\mathbf{x}_k^n = E_{\mathcal{N}\left(\mathbf{x}_k^n | \boldsymbol{\mu}_k^i, \Sigma_k^i\right)}\left[\mathcal{N}\left(\mathbf{f}_k(\mathbf{x}_k^n, \mathbf{u}_k^n), Q_k\right)\right] \tag{41}$$

where

$$\mathbf{f}_k(\mathbf{x}_k^n, \mathbf{u}_k^n) = \begin{bmatrix} \mathbf{f}_k^l(\mathbf{x}_k^n, \mathbf{u}_k^n) \\ \mathbf{f}_k^n(\mathbf{x}_k^n, \mathbf{u}_k^n) \end{bmatrix}$$

Thus, the marginalized Gaussian sum filter decreases the dimensionality of the optimization problem in Ref. [2] but involves an extra expectation type integral in Eq. (41) involving Gaussian pdfs, which can be evaluated using sigma points from the UT. However, the marginalized Gaussian sum approximation may yield more accurate results when compared to the full-state Gaussian sum approximation using the same number of Gaussian components for the nonlinear portion of the state-space model.[b] The increase in accuracy is due to the fact that, in the marginalized approach, the same number of Gaussian components are used to approximate a lower-dimensional pdf.

[b]Based on the assumption that increasing the number of Gaussian components will improve the approximation.

Simulation Results

As an example, consider a nonlinear free fall and parachute model. An 80-kg paratrooper is dropped from an airplane at a height of 1200 m. After 5 seconds, the chute opens. The paratrooper's height as a function of time, $y(t)$, is given by

$$\begin{aligned} \ddot{y}(t) &= -g + \alpha(t)/m \\ y(0) &= 1200 \text{ m} \\ \dot{y}(0) &= 0 \text{ m/s} \end{aligned} \tag{42}$$

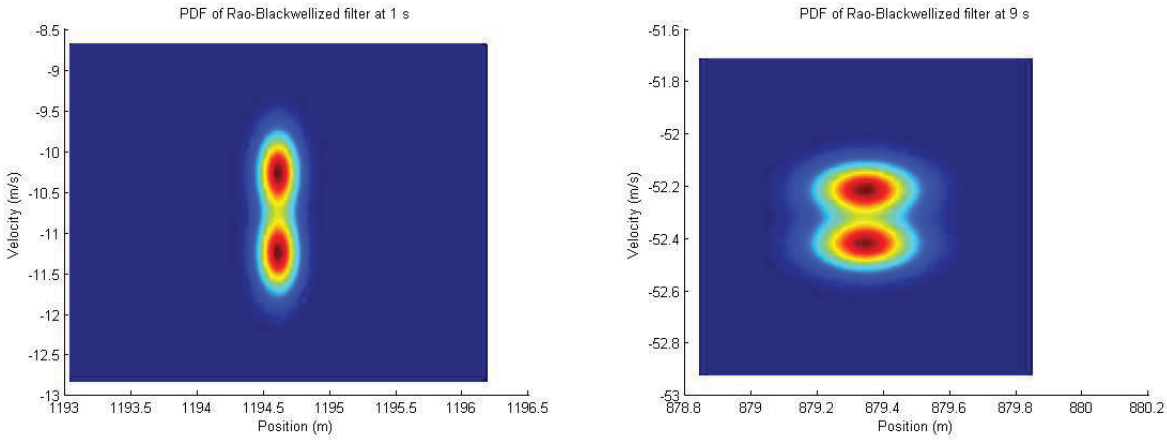
where $g = 9.81$ m/s^2 is the acceleration due to gravity and $m = 80$ kg is the paratrooper's mass. The air resistance $\alpha(t)$ is proportional to the square of the velocity, with different proportionality constants before and after the chute opens:

$$\alpha(t) = \begin{cases} K_1\, \dot{y}(t)^2, & t < 5 \text{ s} \\ K_2\, \dot{y}(t)^2, & t \geq 5 \text{ s} \end{cases} \tag{43}$$

where $K_1 = 1/15$ and $K_2 = 4/15$.

A Monte Carlo simulation is run using 10,000 different initial conditions for a time duration of 15 seconds, with a $\Delta t = 0.1$ seconds. The truth values are computed by taking the average of the Monte Carlo (MC) data points at each time step for the nonlinear model propagation. The simulations for the RB-AGS filter are run 1,000 times using random initializations for the Gaussian components used. Next the average is computed for both methods and used as the overall MC estimated state values and covariances. Lastly, the errors are given by the difference between the MC truth and MC estimate.

In Figures 1(a) and 1(b), the evolution of the pdf of the RB-AGS is shown which clearly indicates the non-Gaussian nature of the problem. In Figure 2 the propagated estimate from the two filters are plotted for the last few iterations. The top plot shows the position estimates, which are both close to the MC solution at each time step. However, the bottom plot reveals the RB-AGS estimate follows the MC values closer than the standard AGS as expected. The resulting estimates only differ by approximately a few tenths of a meter per second between each other, but still obtains a positive improvement in accuracy.



(a) **PDF of RB-AGS Filter After 1 Second** (b) **PDF of RB-AGS Filter After 9 Seconds**

Figure 1. PDF Evolution of RB-AGS Filter

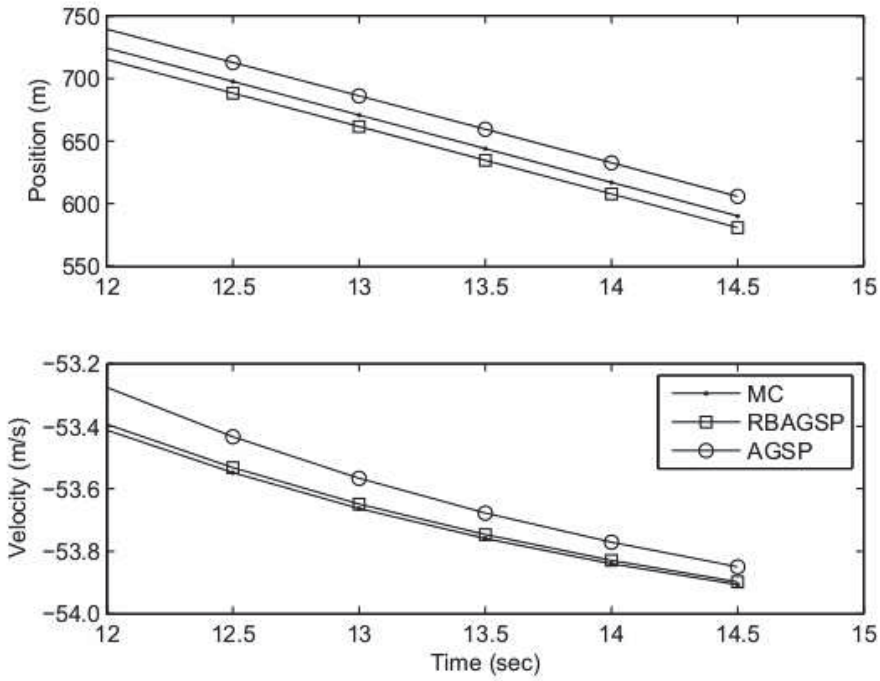


Figure 2. Comparison of the MC Runs and the Propagation Estimates

Conclusions

A marginalized adaptive Gaussian sum filter was presented here. With this new method the linear portion of the state-space model is propagated using the linear Kalman filter and the nonlinear portion using the Unscented Kalman Filter for each Gaussian component. Reducing the linear portion of the state-space model to linear propagation equations places the computational efforts on the nonlinear equations. Simulation results involving parachute model indicates that more accurate results are obtained using the marginalized approach versus the non-marginalized approach. The computational burden using the marginalized approach may be higher or lower than the non-marginalized approach, which is problem dependent.

References

[1]Ito, K. and Xiong, K., "Gaussian Filters for Nonlinear Filtering Problems," *IEEE Transactions on Automatic Control*, Vol. 45, No. 5, 2000, pp. 910–927, doi:10.1109/9.855552.

[2]Terejanu, G., Singla, P., Singh, T., and Scott, P. D., "Uncertainty Propagation for Nonlinear Dynamic Systems Using Gaussian Mixture Models," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 6, Nov.-Dec. 2008, pp. 1623–1633, doi:10.2514/1.36247.

[3]Andrieu, C. and Doucet, A., "Particle Filtering for Partially Observed Gaussian State Space Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 64, No. 4, 2002, pp. 827–836, doi:10.1111/1467-9868.00363.

[4]Gordon, N., Salmond, D., and Smith, A. F. M., "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *Radar and Signal Processing, IEE Proceedings F*, Vol. 140, No. 2, April 1993, pp. 107–103, doi:10.1049/ip-f-2.1993.0015.

[5]Gustafsson, F., Gunnarsson, F., Bergman, N., Forssell, U., Jansson, J., Karlsson, R., and Nordlund, P.-J., "Particle filters for Positioning, Navigation, and Tracking," *IEEE Transactions on Signal Processing*, Vol. 50, No. 2, 2002, pp. 425–437, doi:10.1109/78.978396.

[6]Casella, G. and Robert, C. P., "Rao-Blackwellisation of Sampling Schemes," *Biometrika*, Vol. 83, No. 1, 1996, pp. 81–94, doi:10.1093/biomet/83.1.81.

[7]Schon, T., Gustafsson, F., and Nordlund, P.-J., "Marginalized Particle Filters for Mixed Linear Nonlinear State-Space Models," *IEEE Transactions on Signal Processing*, Vol. 53, 2005, pp. 2279–2289, doi:10.1109/TSP.2005.849151.

[8]Ristic, B., Arulampalam, S., and Gordon, N., *Beyond the Kalman Filter Particle Filters for Tracking Applications*, Artech House, Boston, MA, 2004, p. 5.

[9]Crassidis, J. L. and Junkins, J. L., *Optimal Estimation of Dynamic Systems*, Chapman & Hall/CRC, Boca Raton, FL, 2004, pp. 273–296.

[10]Quiñonero-Candela, J. and Rasmussen, C. E., "Analysis of Some Methods for Reduced Rank Gaussian Process Regression," *Switching and Learning in Feedback Systems, Lecture Notes in Computer Science*, Vol. 3355, Springer-Verlag, Berlin, Germany, Sept. 2005, pp. 98–127, doi:10.1007/978-3-540-30560-6_4.

[11]Rapoport, I. and Oshman, Y., "Fault-Tolerant Particle Filtering by Using Interacting Multiple Model-

Based Rao-Blackwellization," *Journal of Guidance, Control, and Dynamics*, Vol. 28, No. 6, Nov.-Dec. 2005, pp. 1171–1177, doi:10.2514/1.11363.

[12]Schon, T., Karlsson, R., and Gustafsson, F., "The Marginalized Particle Filter in Practice," *2006 IEEE Aerospace Conference*, March 2006, IEEEAC Paper #1671, doi:10.1109/AERO.2006.1655922.