Preprints of the 10th IFAC International Symposium on Dynamics and Control of Process Systems
The International Federation of Automatic Control
December 18-20, 2013. Mumbai, India

Nonlinear Update Based Unscented Gaussian Sum Filter

Krishna Kumar Kottakki, Mani Bhushan, Sharad Bhartiya

Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India - 400076
(E-mail: krishnadfs@iitb.ac.in, mbhushan@che.iitb.ac.in, bhartiya@che.iitb.ac.in)

Abstract: Unscented Kalman Filter (UKF) is a popular method for state and parameter estimation of nonlinear dynamic systems. An attractive feature of UKF is that it utilizes deterministically chosen points (called sigma points), and the number of such points depends linearly on the dimension of the state space. However, an implicit assumption in UKF is that the underlying probability densities are Gaussian. To mitigate the Gaussianity assumption, Gaussian Sum-UKF has been proposed in literature that approximates all underlying densities using a sum of Gaussians. For accurate approximation, the number of sigma points required in this approach is significantly higher than UKF, thereby making the Gaussian Sum-UKF computationally intensive. In this work, we propose an alternate approach labeled unscented Gaussian Sum Filter (UGSF) that leverages the ability of Sum of Gaussians to approximate an arbitrary density, while using the same number of sigma points as in UKF. This is achieved by making suitable design choices of the various parameters in the Gaussian Sum representation. Thus, our approach requires similar computational effort as in UKF and hence does not suffer from the curse of dimensionality. We implement the proposed approach on a nonlinear state estimation case study and demonstrate its superior performance over UKF.

Keywords: Sum of Gaussians, Bayesian estimation, unscented Kalman filter, nonlinear state estimation

1. INTRODUCTION

Recursive Bayesian estimation has been widely used for state and parameter estimation of dynamic systems. Bayesian estimation techniques make use of prior knowledge, typically in the form of a mathematical model of the system, and available measurements to obtain the probability density function of the estimate (Jazwinski, 1970). It can be shown that the conditional posterior mean of the density minimizes the mean square error of the estimator (Mendel, 1995). For linear dynamic systems with linear measurement functions and Gaussian uncertainties in the states and measurements, the posterior density turns out to be Gaussian (Maybeck, 1979). The mean and the covariance of the estimates then correspond to the well-known Kalman filtering linear update equations.

The Kalman filter (KF) for the linear Gaussian system consists of two distinct steps: (i) prediction step: determination of the moments (mean, covariance) of the prior density by propagating through the process model, and (ii) update step: computation of the mean and covariance of the posterior density using a linear update equation (Kalman, 1960).

For nonlinear dynamical systems, analytical solutions to the Bayesian estimation problem are generally not available. Several extensions of the basic Kalman filter have been proposed in literature to address such cases (Patwardhan et al., 2012). Extended Kalman filter (EKF) (Soderstrom, 2003), is one of the most popular extension. The prediction step in EKF involves linearization of system dynamics, thereby leading to a Gaussian prior. Similar to the KF, the update step in EKF is also a linear update step involving linearization of the nonlinear measurement equation. In the last decade, Unscented Kalman filter (UKF) (Julier and Uhlmann, 2004) has emerged as a popular alternate solution to EKF for state estimation of nonlinear dynamical systems (Romanenko and Castro, 2004; Qu and Hahn, 2009; Jacob and Dhib, 2011). It is based on the premise that *"approximating a density function is easier than approximating a function"* (Julier et al., 1995). Similar to EKF, UKF also implicitly assumes the prior density to be Gaussian. However, the mean and covariance of this density are obtained by propagating a set of deterministically chosen points (called sigma points) through the nonlinear process model. This results in better accuracy in the moments of the resulting prior and also avoids numerical issues related to the linearization step of EKF. However, as in EKF, the update equation in UKF is linear with the Kalman gain computed using the predicted sigma points instead of explicit linearization. A key implicit assumption in UKF is that the prior and posterior densities are Gaussian.

Several modifications have been proposed in literature to account for the non-Gaussianity of the various densities.

⋆ The authors thank the Department of Science and Technology, India, for partial financial assistance under grant 09DST010.

These include Gaussian Sum Filter (GSF) (Sorenson and Alspach, 1971; Soderstrom, 2003), Gaussian Sum UKF (Straka et al., 2011) and Sigma point Gaussian sum filter (Šimandl, 2005). These approaches are based on the result that a Gaussian sum can approximate any arbitrary density to any degree of accuracy (Sorenson and Alspach, 1971). Thus, the initial ($k = 0$) posterior is represented as a sum of Gaussians. The prediction and update steps are applied individually to the moments of each Gaussian. At any given instant, the resulting overall prior or posterior density is a weighted sum of the individual Gaussian densities. These weights remain unchanged during the prediction step but are corrected during the update step. The weight-update relies on the innovations and is nonlinear in the measurements. In GSF, the individual prediction and update steps are same as those in EKF (Soderstrom, 2003) and hence suffer from similar issues as in EKF. On the other hand, in Gaussian Sum UKF (GS-UKF) the individual prediction and update steps are same as those in UKF, thereby offering similar advantages as in UKF. An issue in the Gaussian Sum approaches is the appropriate selection of the number of Gaussians to be used. Since the GS-UKF uses the UKF steps to obtain moments separately for each of the individual Gaussians, the total number of sigma points required can be excessively large, thereby leading to a significant computational burden. Moreover, some of the updated weights may become negligible, resulting in a degenerate representation of the prior. Here, only those Gaussians with non-negligible weights contribute to the overall moments. The class of particle filters represents another approach, which does not assume any parameterization of the underlying densities and overcome the issue of degeneracy by an additional resampling step (Arulampalam et al., 2002; Kotecha and Djuric, 2003). However, since these are completely sample based, they require large number of samples and hence are computationally intensive. A detailed survey on nonlinear state estimation techniques can be found in Patwardhan et al. (2012).

In this work, we propose a computationally efficient algorithm for state estimation of nonlinear dynamic systems. The proposed approach utilizes the sigma point concept as used in UKF in the propagation step to obtain the prior density. Subsequently, the propagated sigma points are used to approximate the prior by a sum of Gaussians. The update step is similar to Gaussian Sum filter where the weights are updated nonlinearly with respect to innovations. The proposed approach thus leverages on the computational benefits of UKF and also has the ability to handle non-Gaussian densities as in Gaussian Sum filter. In particular, the prior used in the Bayes' rule is modeled as sum of Gaussians that is obtained based on propagated sigma points chosen deterministically as in UKF. We thus label our approach as Unscented Gaussian Sum Filter (UGSF).

The number of Gaussians in the Gaussian Sum representation of the prior is chosen to be same as the number of sigma points used in the basic UKF. Additionally, the moments of these individual Gaussians in the Gaussian sum representation are chosen such that the moments of the overall density match the UKF moments. These design choices ensure that the computational requirements of our approach are same as UKF while relaxing the implicit assumption inherent in UKF that the prior density is a Gaussian density.

The rest of the paper is organized as follows: The problem statement and the proposed UGSF approach are presented in Section 2. Since, our approach utilizes deterministically chosen sigma points as in UKF, we also briefly summarize UKF and compare our approach with it in Section 2. Section 3 presents a nonlinear state estimation case study to demonstrate the utility of our approach. The paper is then concluded in Section 4.

2. PROBLEM STATEMENT

Consider a discrete-time process and measurement model as,

$$x_{k+1} = f(x_k, u_k) + w_k \qquad w_k \sim \mathcal{N}(0, Q) \tag{1}$$

$$y_k = Hx_k + v_k \qquad v_k \sim \mathcal{N}(0, R) \tag{2}$$

where $f(\cdot)$ is the state dynamics and H is the linear dynamics for observation. $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $u \in \mathbb{R}^p$, $w \in \mathbb{R}^n$, $v \in \mathbb{R}^m$ represent the states, observations, inputs, state noise, and measurement noise, respectively. $Q_{n\times n}$ and $R_{m\times m}$ represent the covariance matrices associated with state and measurement noises, respectively. Note that a linear observation transformation is assumed in Eq (2). This assumption results in a Gaussian likelihood and is made to simplify the posterior density in the proposed UGSF approach. We now summarize the widely used UKF approach for nonlinear state estimation.

2.1 UKF Algorithm (Julier et al., 1995)

At $(k-1)^{th}$ time instant, the sigma points are generated from the given posterior mean $\hat{x}_{k-1|k-1}$ and covariance $P_{k-1|k-1}$.

$$\begin{aligned}
\chi^{(i)}_{k-1|k-1} &= \hat{x}_{k-1|k-1}, \quad i = 0 \\
\chi^{(i)}_{k-1|k-1} &= \hat{x}_{k-1|k-1} + \left[\sqrt{(n+\kappa)P_{k-1|k-1}}\right]_i, \quad i = 1 \dots n \\
\chi^{(i+n)}_{k-1|k-1} &= \hat{x}_{k-1|k-1} - \left[\sqrt{(n+\kappa)P_{k-1|k-1}}\right]_i, \quad i = 1 \dots n
\end{aligned} \tag{3}$$

$$w^{(i)} = \begin{cases} \dfrac{\kappa}{n+\kappa}, & i = 0 \\ \dfrac{1}{2(n+\kappa)}, & i = 1 \dots 2n \end{cases} \tag{4}$$

In the above, $\chi^{(i)}_{k-1|k-1} \in \mathbb{R}^n$ is i^{th} sigma point, $w^{(i)}$ is the weight associated with i^{th} sigma point and κ is a tuning parameter usually chosen as $n + \kappa = 3$.

Prediction Step

These sigma points $\chi^{(i)}_{k-1|k-1}$ are propagated through the process model (Eq (1)) to obtain the predicted sigma points $\chi^{(i)}_{k|k-1}$. The predicted sigma points are subsequently used to construct the prior, which is implicitly assumed to be Gaussian with mean,

$$\hat{x}_{k|k-1} = \sum_{i=0}^{2n} w^{(i)} \chi^{(i)}_{k|k-1} \tag{5}$$

and covariance,

$$P_{k|k-1} = \sum_{i=0}^{2n} w^{(i)}[\chi^{(i)}_{k|k-1} - \hat{x}_{k|k-1}][\chi^{(i)}_{k|k-1} - \hat{x}_{k|k-1}]^T + Q \tag{6}$$

The following quantities are then computed to evaluate the Kalman gain,

$$\eta^{(i)}_{k|k-1} = H\chi^{(i)}_{k|k-1} \tag{7}$$

$$\hat{\eta}_{k|k-1} = \sum_{i=0}^{2n} w^{(i)}\eta^{(i)}_{k|k-1} \tag{8}$$

$$P_{xy} = \sum_{i=0}^{2n} w^{(i)}[\chi^{(i)}_{k|k-1} - \hat{x}_{k|k-1}][\eta^{(i)}_{k|k-1} - \hat{\eta}_{k|k-1}]^T + QH^T \tag{9}$$

$$P_{yy} = \sum_{i=0}^{2n} w^{(i)}[\eta^{(i)}_{k|k-1} - \hat{\eta}_{k|k-1}][\eta^{(i)}_{k|k-1} - \hat{\eta}_{k|k-1}]^T + HQH^T + R \tag{10}$$

$$K_k = P_{xy}{P_{yy}}^{-1} \tag{11}$$

In Eqs (7) to (11), $\eta^{(i)}_{k|k-1}$, $\hat{\eta}_{k|k-1} \in \mathbb{R}^m$, $P_{xy} \in \mathbb{R}^{n\times m}$, $P_{yy} \in \mathbb{R}^{m\times m}$ and $K_k \in \mathbb{R}^{n\times m}$ represent transformed sigma points (in measurement space), weighted mean of the transformed sigma points, cross covariance of state estimation error and innovation, innovation covariance, and Kalman Gain, respectively.

Linear Update Step

Moments associated with $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ are then updated linearly as,

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{\eta}_{k|k-1}) \tag{12}$$

$$P_{k|k} = P_{k|k-1} - K_kP_{k|k-1}K_k^T \tag{13}$$

Here, $\hat{x}_{k|k}$, $P_{k|k}$ are updated mean and updated covariance respectively. This completes one cycle of UKF. Next we present the proposed UGSF approach.

2.2 The Unscented Gaussian Sum Filter- Proposed approach

A nonlinear transformation does not preserve Gaussianity in general (Sorenson and Alspach, 1971). UKF overlooks this underlying fact as given in Eqs (5) and (6). To overcome this problem, the prior density can be approximated as a sum of Gaussians as discussed in Section 1. We now discuss our proposed approach for obtaining this prior approximation followed by a discussion of the update step to obtain the posterior.

(i) Prior Approximation with sum of Gaussians

We now approximate the prior with sum of Gaussians as,

$$p_{x_k|Y_{k-1}}(\xi_k|Y_{k-1}) = \sum_{i=1}^{N} \propto^{(i)} p^{(i)}(\xi_k) \tag{14}$$

$$p^{(i)} \sim \mathcal{N}(\mu^{(i)}_{k|k-1}, \Sigma^{(i)}_{k|k-1}), \quad i = 1, 2, \dots N \tag{15}$$

In Eqs (14) and (15), $p_{x_k|Y_{k-1}}(\xi_k|Y_{k-1})$ is prior density of the states, $\xi_k \in \mathbb{R}^n$ is the dummy variable associated with the state x_k, and Y_{k-1} is a set of all previous observations i.e $\{y_1, y_2 \dots y_{k-1}\}$. $p^{(i)}$ is a Gaussian density with mean $\mu^{(i)}_{k|k-1}$ and covariance $\Sigma^{(i)}_{k|k-1}$, and $\propto^{(i)}$ is the non-negative weight associated with $p^{(i)}$ such that $0 <\propto^{(i)}< 1$ and $\sum_{i=1}^{N} \propto^{(i)}= 1$. Further, N is the number of individual Gaussians in the Gaussian sum representation. With appropriate choice of parameters N, $\propto^{(i)}$, $\mu^{(i)}_{k|k-1}$ and $\Sigma^{(i)}_{k|k-1}$ any prior density can be approximated by Eq (14) to an arbitrary degree of accuracy. A logical question that arises is how to select the number of Gaussians as well as the parameters in this Gaussian sum representation. In this work, we use the unscented transformation idea as practiced in UKF to achieve this representation. In particular, we make the following design choices for these parameters,

- $N = 2n + 1$; number of Gaussians is same as number of sigma points,
- $\mu^{(i)}_{k|k-1} = \chi^{(i)}_{k|k-1}, i = 1 \dots N$; each i^{th} Gaussian is assigned i^{th} propagated sigma point as its mean,
- $\propto^{(i)}= w^{(i)}$; each i^{th} Gaussian is assigned the weight corresponding to the i^{th} sigma point,
- $\Sigma^{(i)}_{k|k-1} = Q$; the covariance of each individual Gaussian is the same as the covariance of process noise.

To be consistent with the UKF notation, the individual Gaussians are numbered from 0 to $N-1$ i.e from 0 to $2n$. The above choice of the parameters ensures that the mean and covariance of proposed sum of Gaussian prior matches with the single Gaussian prior implicitly assumed in UKF as shown below.

Prior Mean

For a density represented as a sum of Gaussians as in Eq (14), the first moment can be obtained as,

$$E[x_k|Y_{k-1}] \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} \xi_k p_{x_k|Y_{k-1}}(\xi_k|Y_{k-1})d\xi_k = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} \xi_k \sum_{i=1}^{N} \propto^{(i)} p^{(i)}(\xi_k)d\xi_k \tag{16}$$

Using the design choices outlined above, the first moment in the proposed approach turns out to be,

$$E[x_k|Y_{k-1}] = \sum_{i=0}^{2n} w^{(i)}\chi^{(i)}_{k|k-1} = \hat{x}_{k|k-1} \tag{17}$$

Prior Covariance

The covariance of the Gaussian sum prior can be expressed in terms of the moments of the individual Gaussians (Soderstrom, 2003),

$$P_{k|k-1} = \sum_{i=1}^{N} \alpha^{(i)}[\Sigma^{(i)}_{k|k-1} + \{\cdot\}\{\cdot\}^T] \tag{18}$$

$$\text{where, } \{\cdot\} = [\mu^{(i)}_{k|k-1} - E[x_k|Y_{k-1}]] \tag{19}$$

Upon implementing the design choices discussed above, the covariance simplifies to,

$$P_{k|k-1} = \sum_{i=0}^{2n} w^{(i)}[\chi^{(i)}_{k|k-1} - \hat{x}_{k|k-1}][\chi^{(i)}_{k|k-1} - \hat{x}_{k|k-1}]^T + Q \tag{20}$$

which is identical to the prior covariance obtained in the UKF approach (Eq (6)). Thus, the Gaussian sum prior density used in UGSF and represented in Eq (14) becomes,

$$p_{x_k|Y_{k-1}}(\xi_k|Y_{k-1}) = \sum_{i=0}^{2n} w^{(i)} \frac{1}{(2\pi)^{n/2}|Q|^{1/2}} \exp\{\cdot\} \tag{21}$$

$$\text{where, } \{\cdot\} = \left\{ \frac{-1}{2}[\xi_k - \chi^{(i)}_{k|k-1}]^T Q^{-1}[\xi_k - \chi^{(i)}_{k|k-1}] \right\} \tag{22}$$

(ii) Update of the Gaussian sum prior

The Bayes' rule for obtaining the posterior density from the prior is given as (Maybeck, 1979),

$$p_{x_k|Y_k}(\xi_k|Y_k) = \frac{p_{x_k|Y_{k-1}}(\xi_k|Y_{k-1}) p_{y_k|x_k,Y_{k-1}}(\zeta_k|\xi_k, Y_{k-1})}{p_{y_k|Y_{k-1}}(\zeta_k|Y_{k-1})} \tag{23}$$

In Eq (23), the prior density is $p_{x_k|Y_{k-1}}(\xi_k|Y_{k-1})$ described in Eq (21). Further, the likelihood density in Eq (23) $p_{y_k|x_k,Y_{k-1}}(\zeta_k|\xi_k, Y_{k-1})$ takes into account the stochastic component associated with the observation model in Eq (2), which is assumed to be Gaussian and thus turns out to be (Maybeck, 1979),

$$p_{y_k|x_k,Y_{k-1}}(\zeta_k|\xi_k, Y_{k-1}) = \frac{1}{(2\pi)^{m/2}|R|^{1/2}} \exp\{\cdot\} \tag{24}$$

$$\text{where, } \{\cdot\} = \left\{ \frac{-1}{2}[y_k - H\xi_k]^T R^{-1}[y_k - H\xi_k] \right\} \tag{25}$$

$p_{y_k|Y_{k-1}}(\zeta_k|Y_{k-1})$ is a normalization factor called evidence and is defined as (Maybeck, 1979),

$$p_{y_k|Y_{k-1}}(\zeta_k|Y_{k-1}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \mathbf{p}(\xi_k) d\xi_k \tag{26}$$

where,

$$\mathbf{p}(\xi_k) = p_{x_k|Y_{k-1}}(\xi_k|Y_{k-1}) p_{y_k|x_k,Y_{k-1}}(\zeta_k|\xi_k, Y_{k-1}) \tag{27}$$

Substituting Eqs (21) and (24) into (26) and integrating yields,

$$p_{y_k|Y_{k-1}}(\zeta_k|Y_{k-1}) = \propto (2\pi)^{n/2}|\Xi|^{1/2} \sum_{i=0}^{2n} w^{(i)} \exp(-\frac{1}{2}\epsilon_k^{(i)}) \tag{28}$$

where,

$$\propto = \frac{1}{((2\pi)^{(m+n)}|Q||R|)^{1/2}} \tag{29}$$

$$\Xi = [Q^{-1} + H^T R^{-1} H]^{-1} \tag{30}$$

$$\epsilon_k^{(i)} = (y_k - H\chi^{(i)}_{k|k-1})^T [HQH^T + R]^{-1} (y_k - H\chi^{(i)}_{k|k-1}) \tag{31}$$

In Eqs (24) and (28), the dummy variable ζ_k on the right hand side is substituted by its realization namely, y_k. The analytical solution for the posterior density can now be obtained by substituting Eqs (21), (24) and (28) in Eq (23). This results in the following posterior density,

$$p_{x_k|Y_k}(\xi_k|Y_k) = \sum_{i=0}^{2n} \frac{\delta_k^{(i)}}{(2\pi)^{n/2}|\Xi|^{1/2}} \exp\{\cdot\} \tag{32}$$

where,

$$\{\cdot\} = \left\{ \frac{-1}{2}(\xi_k - \chi^{(i)}_{k|k})^T \Xi^{-1} (\xi_k - \chi^{(i)}_{k|k}) \right\} \tag{33}$$

$$\chi^{(i)}_{k|k} = \Xi[Q^{-1}\chi^{(i)}_{k|k-1} + H^T R^{-1}\mathbf{y}_k] \tag{34}$$

$$\delta_k^{(i)} = \frac{\tilde{w}_k^{(i)}}{\sum_{j=0}^{2n} \tilde{w}_k^{(j)}} \tag{35}$$

In Eq (32), the posterior density is a sum of Gaussians with each Gaussian being centered around the mean $\chi^{(i)}_{k|k}$, covariance Ξ and weighted with $\delta_k^{(i)}$. These weights are nonlinear in the individual innovations and are identical to those obtained in Gaussian Sum filters (Soderstrom, 2003). The updated mean and updated covariance can now be obtained from the posterior density analytically. The posterior mean turns out to be,

$$\hat{x}_{k|k} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \xi_k p_{\mathbf{x}_k|\mathbf{Y}_k}(\xi_k|Y_k) d\xi_k = \sum_{i=0}^{2n} \delta_k^{(i)} \chi^{(i)}_{k|k} \tag{36}$$

The posterior covariance turns out to be,

$$P_{k|k} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} (\xi_k - \hat{x}_{k|k})(\xi_k - \hat{x}_{k|k})^T p_{x_k|y_k}(\xi_k|Y_k) d\xi_k \tag{37}$$

$$= \Xi + \sum_{i=0}^{2n} \delta_k^{(i)} [\chi^{(i)}_{k|k} - \hat{x}_{k|k}][\chi^{(i)}_{k|k} - \hat{x}_{k|k}]^T \tag{38}$$

In our proposed approach, the sigma points at instant k required for the next cycle are obtained from a Gaussian with the mean and covariance being the same as that of the Gaussian Sum posterior given above (Eqs (36) and (38)). We now summarize the proposed UGSF approach as an online state estimator.

UGSF Summary

Step 1: Generate $2n+1$ deterministic samples $\chi^{(i)}_{k-1|k-1}$ and their corresponding weights $w^{(i)}$ from the initial prior density $\mathcal{N}(\hat{x}_{k-1|k-1}, P_{k-1|k-1})$ using Eq (3).

Step 2: Propagate each sigma point, $\chi^{(i)}_{k-1|k-1}$ through nonlinear process dynamics (Eq (1)) to obtain the predicted sigma points $\chi^{(i)}_{k|k-1}$.

Step 3: Use the analytical expressions from Eqs (36) and (37), to evaluate mean and covariance of posterior density at k^{th} instant.

Step 4: One cycle is completed now. Go to Step (1) with $k = k+1$.

We now discuss the key differences between the proposed approach and the existing GS-UKF (Straka et al., 2011).

(1) Approximation of Prior: In the proposed approach, the number of Gaussians in the Gaussian sum representation of the prior density at any time instant k is chosen as $2n+1$, which is the same as the number of sigma points typically generated in UKF. In the GS-UKF approach on the other hand, the number of Gaussians in the initial posterior representation (at $k = 0$) is a user-specified parameter, typically

chosen to be large to obtain a good approximation. Another difference in the two approaches is in the computation of the mean and covariance of the individual Gaussians in the Gaussian sum representation. In the proposed UGSF approach, covariance of each individual Gaussian is chosen to be Q while its mean is simply the propagated value of the corresponding sigma point. These choices ensure that the overall mean and covariance of the Gaussian sum representation is the same as the mean and covariance that would be obtained if basic UKF had been used. On the other hand, in the existing GS-UKF approach, the mean and covariance of an individual Gaussian is obtained by propagating sigma points generated locally for each Gaussian (Straka et al., 2011). Hence, the GS-UKF prior mean and covariance do not match with the UKF moments.

(2) Approximation of Posterior: As seen in Section 2.2 earlier, the posterior obtained by Bayes' rule turns out to be a sum of Gaussians (Eq (32)) when the prior is represented as a sum of Gaussians. In the proposed UGSF approach, this posterior is reapproximated as a single Gaussian having same mean and covariance as the sum of Gaussians (Eqs (36) and (38)). In the existing GS-UKF approach, this re-approximation is not performed and hence can lead to an increase in the number of Gaussians in subsequent steps. Pruning strategies have to be used to limit this number to an acceptable value. Further, absence of re-approximation can potentially lead to the degeneracy issues in GS-UKF as discussed previously.

(3) Computational Issues: In the state-estimation filters relying on Unscented Transformation, such as UKF, GS-UKF, UGSF, the main computational effort is in propagating the sigma points through the nonlinear process model. Since, the number of sigma points in our proposed UGSF approach are same as that in UKF, we expect the computational requirements of the two to be similar. The existing GS-UKF approach requires propagation of $N(2n+1)$ sigma points where N is the number of individual Gaussians in the posterior representation, typically larger than $2n+1$. This has the potential to obtain a better representation of the underlying densities than UGSF, but at a significantly higher computational cost.

3. RESULTS AND DISCUSSIONS

The proposed UGSF and UKF are now implemented on a two state non-isothermal CSTR problem (Bruns and Bailey, 1977). The average performances of both UKF and UGSF filters are compared for 100 different noise realizations over 1000 time instants, each, for different values of Q and R.

3.1 Non-Isothermal CSTR

Dynamics of the CSTR in dimensionless form is given as (Bruns and Bailey, 1977),

$$\begin{aligned} \frac{dx_1}{dt} &= 1 - x_1 - Dax_1 e^{\gamma(1-\frac{1}{x_2})} + w_1 \\ \frac{dx_2}{dt} &= 1 - x_2 + Da\beta x_1 e^{\gamma(1-\frac{1}{x_2})} + \vartheta(u - x_2) + w_2 \\ y &= x_2 + v \end{aligned} \tag{39}$$

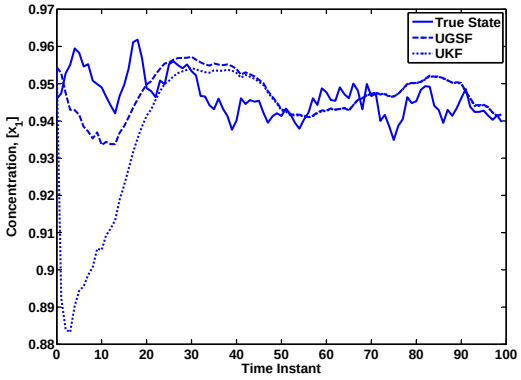

Fig. 1. States tracking of UKF and UGSF for x_1

where, x_1, x_2, u and t are the dimensionless reactant concentration, reactor temperature, coolant temperature, and time, respectively. w_1, w_2 are the process noises with covariance Q and v is the measurement noise with covariance R. Base values of state and measurement noise standard deviations correspond to 1% of the corresponding steady state values. Further the process input u is perturbed as a PRBS signal (MATLAB®) around the steady state value u_s with standard deviation being 2.5% of u_s. The various process parameters are listed in Table 1. Performance of

Table 1. Process parameters for CSTR case study (Bruns and Bailey, 1977)

$Da = 0.0637$	$x_{ss} = [0.94500 \quad 0.9976]^T$ (steady state)
$\beta = 0.167$	$x_{0\|0} = [0.95445 \quad 1.0075]^T$
$\gamma = 37.7$	$Q_{base} = diag\{9.025 \times 10^{-5}, 9.94 \times 10^{-5}\}$
$\vartheta = 0.1356$	$R_{base} = 9.94 \times 10^{-5}$
$u_s = 1.048$	$T_s = 0.1t$
$\kappa = 1$	$P_{0\|0} = diag\{9.025 \times 10^{-3}, 9.94 \times 10^{-3}\}$
	$\hat{x}_{0\|0} = 1.01 \times x_{ss}$

both UKF and UGSF are compared based on values of Average Root Mean Square Error (ARMSE) defined as,

$$e_i = \frac{1}{K}\sum_{l=1}^{K}\left[\sqrt{\frac{1}{M}\sum_{j=1}^{M}[\hat{x}_{ji}^{(l)} - x_{ji}^{(l)}]^2}\right] \quad i = 1, \ldots, n \tag{40}$$

where, K, M and $x_{ji}^{(l)}$ represent number of noise realizations, number of time instances and the i^{th} state at j^{th} time instant for l^{th} noise realization. The results

Table 2. ARMSE values of UKF and UGSF for CSTR case study

Mode	UKF		UGSF	
	$x_1 \times 10^3$	$x_2 \times 10^3$	$x_1 \times 10^3$	$x_2 \times 10^3$
$Q_1; R_1$	23.0169	7.66092	22.4141	7.65857
$Q_1; R_2$	22.4790	3.0144	21.8920	3.0143
$Q_1; R_3$	22.4220	0.99166	21.8390	0.99166
$Q_2; R_1$	8.9476	4.4762	7.7708	4.4684
$Q_2; R_2$	8.2880	2.4243	7.0945	2.4270
$Q_2; R_3$	8.2419	0.95325	7.0321	0.9534
$Q_3; R_1$	5.2542	1.9174	3.0993	1.9044
$Q_3; R_2$	4.9435	1.41488	3.6377	1.4437
$Q_3; R_3$	4.8257	0.76650	6.1477	0.7777

are presented in Table 2. In this Table, $Q_1 = Q_{base}$, $Q_2 = 0.1Q_{base}$ and $Q_3 = 0.01Q_{base}$. Similarly, $R_1 = R_{base}$, $R_2 = 0.1R_{base}$ and $R_3 = 0.01R_{base}$. From the results it

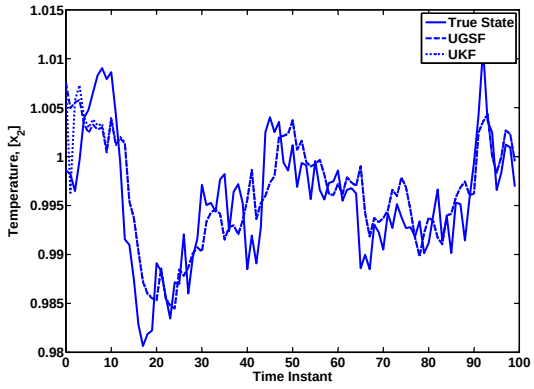

Fig. 2. States tracking of UKF and UGSF for x_2

can be observed that the UGSF performance is superior to the UKF performance for most of the cases considered. However, out of the 9 cases, for three cases for state x_2 $(Q_2, R_2; Q_2, R_3; Q_3, R_2)$ and one case for state x_1 (Q_3, R_3) the performance of UKF was better. These cases where the UGSF performance was inferior typically correspond to low values of Q and R. This inferior performance may be due to inadequate approximation of the prior with the limited number of sigma points (5 in this case). The concentration estimates from UKF and UGSF along with the true states are shown in Fig. 1 while the estimated temperatures (x_2) along with the measurements are shown in Fig. 2. These plots correspond to (Q_2, R_1) combination for one representative simulation run from amongst the 100 different noise realizations. While the actual simulation was carried out for 1000 time instants, the results are shown only up to 100 time instants for visual clarity. The above analysis was carried out for a fixed initial condition as reported in Table 2. To compare the performances for varying initial conditions, we performed state estimation for two combinations of Q and R with 100 different randomly chosen initial conditions. These two combinations were: (1) Q_1, R_2; and (2) Q_2, R_1. The corresponding ARMSE values $(\times 10^3)$ for UKF and UGSF were obtained as: (1) x_1: 21.888 and 21.4583, x_2: 3.0555 and 3.0555, (2) x_1: 7.776 and 7.0388, x_2: 4.3936 and 4.3906. These results once again demonstrate the superior performance of UGSF compared to UKF. Additionally, we also implemented GS-UKF with $N = 5$ for $Q = Q_1$, $R = R_1$. The average computation time per time instant for UKF, UGSF and GS-UKF were found to be 0.0176, 0.0176 and 0.0879 seconds, respectively. Also, the performance of UGSF was found superior to GS-UKF and UKF performances (Case I, Table 2).

4. CONCLUSIONS

This work presents a computationally efficient extension of UKF, labeled UGSF, to exploit the representation ability of sum of Gaussians for nonlinear state estimation. In the proposed approach, the number of Gaussians in the Gaussian sum representation of prior is restricted to the number of sigma points chosen in UKF, thereby ensuring similar computational requirements as UKF. The case study demonstrates that better estimation performance as compared to UKF is obtained by the proposed UGSF approach. Thus, the proposed approach represents a trade-off for nonlinear state estimation, where underlying densities tend to be non-Gaussian, between the potentially more accurate but computationally demanding GS-UKF as existing in literature and the relatively less accurate UKF that requires much lower computations.

REFERENCES

Bruns, D. and Bailey, J. (1977). Nonlinear feedback control for operating a nonisothermal cstr near an unstable steady state. *Chemical Engineering Science*, 32(3), 257 – 264.

Jacob, N.C. and Dhib, R. (2011). Unscented kalman filter based nonlinear model predictive control of a low density poly ethylene autoclave reactor. *Journal of Process Control*, 21(9), 1332 – 1344.

Jazwinski, A. (1970). *Stochastic Process and Filtering Theory*, volume 141 of *Academic Press*.

Julier, S. and Uhlmann, J. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3), 401 – 422.

Julier, S., Uhlmann, J., and Durrant-Whyte, H. (1995). A new approach for filtering nonlinear systems. In *American Control Conference, Proceedings of the 1995*, volume 3, 1628–1632 vol.3.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D), 35–45.

Kotecha, J. and Djuric, P. (2003). Gaussian sum particle filtering. *IEEE Transactions on Signal Processing*, 51(10), 2602 – 2612.

Maybeck, P.S. (1979). *Stochastic models, estimation, and control*, volume 141 of *Mathematics in Science and Engineering*.

Mendel, J. (1995). *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Pearson Education.

Patwardhan, S.C., Narasimhan, S., Jagadeesan, P., Gopaluni, B., and Shah, S.L. (2012). Nonlinear bayesian state estimation: A review of recent developments. *Control Engineering Practice*, 20(10), 933 – 953.

Qu, C.C. and Hahn, J. (2009). Process monitoring and parameter estimation via unscented kalman filtering. *Journal of Loss Prevention in the Process Industries*, 22(6), 703 – 709.

Romanenko, A. and Castro, J.A. (2004). The unscented filter as an alternative to the ekf for nonlinear state estimation: a simulation case study. *Computers and Chemical Engineering*, 28(3), 347 – 355.

Sanjeev Arulampalam, M., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2), 174–188.

Soderstrom, T. (2003). *Discrete time stochastic systems*, volume 141 of *Control and Signal processing*.

Sorenson, H. and Alspach, D. (1971). Recursive bayesian estimation using gaussian sums. *Automatica*, 7(4), 465 – 479.

Straka, O., Dunik, J., and Simandl, M. (2011). Gaussian sum unscented kalman filter with adaptive scaling parameters. In *Proceedings of the 14th International Conference on Information Fusion (FUSION)*, 1 –8.

Šimandl, M.D.J. (2005). Sigma point gaussian sum filter design using square root unscented filters. *Proceedings of the 16th IFAC World Congress*, 16, 1000–1005.