Bayesian Unscented Kalman Filter for State Estimation of Nonlinear and Non-Gaussian Systems

Zhong Liu, Shing-Chow Chan, Ho-Chun Wu and Jiafei Wu
Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, P. R. China
{liuzhong, scchan, andrewhcwu, jfwu}@eee.hku.hk

***Abstract*—This paper proposes a Bayesian unscented Kalman filter with simplified Gaussian mixtures (BUKF-SGM) for dynamic state space estimation of nonlinear and non-Gaussian systems. In the BUKF-SGM, the state and noise densities are approximated as finite Gaussian mixtures, in which the mean and covariance for each component are recursively estimated using the UKF. To avoid the exponential growth of mixture components, a Gaussian mixture simplification algorithm is employed to reduce the number of mixture components, which leads to lower complexity in comparing with conventional re-sampling and clustering techniques. Experimental results show that the proposed BUKF-SGM can achieve better performance compared with the particle filter (PF)-based algorithms. This provides an attractive alternative for nonlinear state estimation problem.**

Keywords—Bayesian unscented Kalman filter; dynamic state estimation; nonlinear and non-Gaussian system; Gaussian mixture; particle filter.

I. INTRODUCTION

State estimation of stochastic dynamic systems from a sequence of noisy observations plays a crucial role in many practical applications such as target tracking, fault detection, signal processing and automatic control problems. The dynamic system under consideration is usually modeled by the state-space approach, where difference or differential equations are used to model the evolution of the system over time.

When the model is linear with Gaussian distributed noises, the celebrated Kalman filter (KF) [1] is an optimal state estimator which can provide the mean and covariance sequentially. For more complicated systems with nonlinear models and non-Gaussian noises, there is no closed form analytic expression for the posteriori densities. Therefore, the state and noise densities have to be approximated by a variety of methods such as histogram or Gaussian mixtures equipped with nonlinear filters say the extended Kalman filter (EKF) and unscented Kalman filter (UKF) [2]. The histogram approach represents the density as a set of particles and it gives rise to the particle filtering technique [3], which approximates the non-Gaussian densities with a set of weighted particles. However the computational complexity of such methods grows exponentially with the dimension of the states and particles. In addition, due to the problems of degeneracy and sampling impoverishment [4], the standard PF may be rather inefficient in sampling. To alleviate the problems, many particle filter (PF)-based algorithms [5-7] have been proposed. They employ re-sampling and clustering techniques to simplify the density to achieve reduction in complexity. For instance, in [6], a greedy expectation maximization (EM)-based algorithm for model order reduction using Kullback-Leibler divergence (KLD) [8] has been employed to simplify the density after re-sampling. Similar concept has been extended to a nonlinear setting as the Gaussian sum particle filter (GSPF) [5], in which the state density for mixture components is approximated using weighted particles.

Recently, for linear non-Gaussian systems, a Bayesian Kalman filter (BKF) has been reported in our previous work [9], which is able to approximate the non-Gaussian noise as simplified Gaussian mixtures with relatively low complexity. The effectiveness of the BKF algorithm has been demonstrated in video object tracking and other applications [10]. Since state estimation of nonlinear systems is frequently encountered in many practical applications, it will be highly desirable if these nice properties can be extended to the nonlinear settings.

It is the aim of this paper to further extend the BKF concept to nonlinear state estimation using the unscented transformation technique. In particular, we propose a new Bayesian unscented Kalman filter with simplified Gaussian mixtures (BUKF-SGM) for state estimation in non-linear and non-Gaussian systems. Among different nonlinear filtering methods, we have chosen the unscented transform as it is renowned for its well performance in approximating the mean and covariance of the state using a minimal set of sample points, rather than linearizing the nonlinear function as in the EKF. While the nonlinearity is captured using the unscented transform, the state density in the proposed approach is modeled using an efficient GM simplification procedure in BKF, which directly simplifies the GMs by minimizing an upper bound of the approximation error between the original and the simplified models. This avoids the high complexity of the re-sampling and clustering encountered in [5-7], while

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

allowing more general state and measurement noises to be modeled resulting in better tracking of the states. Like its BKF counterpart, the arithmetic complexity of the new BUKF-SGM algorithm is only a polynomial, instead of exponential, function of the state dimension. The proposed BKF-SGM is applied to solve a nonlinear and non-Gaussian time series estimation problem and the simulation results show that it is able to achieve a better tracking accuracy than the PF-based algorithms. It therefore serves an attractive alternative to PF for nonlinear state estimation, especially for systems with more number of states. The rest of the paper is organized as follows: Section II describes the details of the proposed BUKF-SGM algorithm. In Section III, the performance of the BUKF-SGM is evaluated and compared with the PF-based algorithms. The conclusion is drawn in Section IV.

II. BUKF-SGM For State Estimation Of Nonlinear and Non-Gaussian Systems

For simplicity, we consider the following autonomous discrete-time nonlinear state-space model:

$$\boldsymbol{x}_k = f_k(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_k, \tag{1}$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k) + \boldsymbol{v}_k, \tag{2}$$

where $\boldsymbol{x}_k$ and $\boldsymbol{z}_k$ denote respectively the state and observation vectors at time instant k, and $f_k(\boldsymbol{x}_{k-1})$ and $h_k(\boldsymbol{x}_k)$ are respectively known nonlinear state and measurement functions. $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are random vectors of given distributions, which are assumed to be mutually independent and independent of $\boldsymbol{x}_k$.

2.1 Unscented Kalman Filter (UKF)

In the UKF framework, the probability density functions (pdfs) of the state noise $\boldsymbol{w}_k$ and observation noise $\boldsymbol{v}_k$ are usually assumed to be zero mean Gaussian distributed, i.e.:

$$p(\boldsymbol{w}_k) = N(\boldsymbol{w}_k, \overline{\boldsymbol{w}}_k, \boldsymbol{Q}_k), \tag{3}$$

$$p(\boldsymbol{v}_k) = N(\boldsymbol{v}_k, \overline{\boldsymbol{v}}_k, \boldsymbol{R}_k), \tag{4}$$

with $\overline{\boldsymbol{w}}_k = 0$ and $\overline{\boldsymbol{v}}_k = 0$, and $N(\boldsymbol{u}_k, \overline{\boldsymbol{u}}_k, \boldsymbol{C}_k)$ denotes a Gaussian distribution with mean $\overline{\boldsymbol{u}}_k$ and covariance $\boldsymbol{C}_k$. Rather than linearizing the nonlinear functions as in the EKF, UKF utilizes an unscented transform to approximate the non-Gaussian state density arising from the nonlinearity by a Gaussian distribution through sampling at a set of sigma points. To be specific, given an L-dimensional state vector $\boldsymbol{x}_k$ with mean $\boldsymbol{\mu}_k$ and covariance $\boldsymbol{P}_k$, the UKF algorithm for state estimation can be summarized as follows:

1) Construct the following sigma points $\{\boldsymbol{\chi}_{l(k-1)}\}$ with corresponding weights $\{w_l\}$ from the Gaussian distribution with mean $\boldsymbol{\mu}_{k-1}$ and covariance $\boldsymbol{P}_{k-1}$:

$$\boldsymbol{\chi}_{0(k-1)} = \boldsymbol{\mu}_{k-1},\ w_0 = \frac{\kappa}{L+\kappa}, \tag{5}$$

$$\boldsymbol{\chi}_{l(k-1)} = \boldsymbol{\mu}_{k-1} + \left(\sqrt{(L+\kappa)\boldsymbol{P}_{k-1}}\right)_l,\ w_l = \frac{1}{2(L+\kappa)}, \tag{6}$$

$$\boldsymbol{\chi}_{(L+l)(k-1)} = \boldsymbol{\mu}_{k-1} - \left(\sqrt{(L+\kappa)\boldsymbol{P}_{k-1}}\right)_l,\ w_{L+l} = w_l, \tag{7}$$

where $l = 1,...,L$, κ is the scaling parameter and $\left(\sqrt{(L+\kappa)\boldsymbol{P}_{k-1}}\right)_l$ represents the l-th column of the matrix $\left(\sqrt{(L+\kappa)\boldsymbol{P}_{k-1}}\right)$.

2) Propagate each sigma point through the nonlinear system state function:

$$\boldsymbol{\chi}_{l(k|k-1)} = f_k(\boldsymbol{\chi}_{l(k-1)}),\quad \forall l. \tag{8}$$

3) Approximate the predicted density after the nonlinear state function by a Gaussian distribution by using the ensemble mean and covariance estimates as follows:

$$\boldsymbol{\mu}_{k|k-1} = \sum_{l=0}^{2L} w_l \boldsymbol{\chi}_{l(k|k-1)}, \tag{9}$$

$$\boldsymbol{P}_{k|k-1} = \sum_{l=0}^{2L} w_l (\boldsymbol{\chi}_{l(k|k-1)} - \boldsymbol{\mu}_{k|k-1})(\boldsymbol{\chi}_{l(k|k-1)} - \boldsymbol{\mu}_{k|k-1})^T + \boldsymbol{Q}_k. \tag{10}$$

4) Construct a similar set of sigma points from the approximately predicted Gaussian distribution with mean $\boldsymbol{\mu}_{k|k-1}$ and covariance $\boldsymbol{P}_{k|k-1}$:

$$\boldsymbol{\chi}_{0(k|k-1)} = \boldsymbol{\mu}_{k|k-1},\ w_0 = \frac{\kappa}{L+\kappa}, \tag{11}$$

$$\boldsymbol{\chi}_{l(k|k-1)} = \boldsymbol{\mu}_{k|k-1} + \left(\sqrt{(L+\kappa)\boldsymbol{P}_{k|k-1}}\right)_l,\ w_l = \frac{1}{2(L+\kappa)}, \tag{12}$$

$$\boldsymbol{\chi}_{(L+l)(k|k-1)} = \boldsymbol{\mu}_{k|k-1} - \left(\sqrt{(L+\kappa)\boldsymbol{P}_{k|k-1}}\right)_l,\ w_{L+l} = w_l, \tag{13}$$

where $l = 1,...,L$.

5) Propagate the sigma points through the nonlinear system measurement function and calculate the predicted measurement mean:

$$\boldsymbol{z}_{l(k|k-1)} = h_k(\boldsymbol{\chi}_{l(k|k-1)}),\quad \forall l \tag{14}$$

$$\boldsymbol{z}_{k|k-1} = \sum_{l=0}^{2L} w_l \boldsymbol{z}_{l(k|k-1)}. \tag{15}$$

6) Estimate the measurement covariance and cross-covariance:

$$\boldsymbol{P}_{z,k|k-1} = \sum_{l=0}^{2L} w_l (\boldsymbol{z}_{l(k|k-1)} - \boldsymbol{z}_{k|k-1})(\boldsymbol{z}_{l(k|k-1)} - \boldsymbol{z}_{k|k-1})^T + \boldsymbol{R}_k, \tag{16}$$

$$\boldsymbol{P}_{xz,k|k-1} = \sum_{l=0}^{2L} w_l (\boldsymbol{\chi}_{l(k|k-1)} - \boldsymbol{\mu}_{k|k-1})(\boldsymbol{z}_{l(k|k-1)} - \boldsymbol{z}_{k|k-1})^T. \tag{17}$$

7) Calculate the posteriori estimation of the state and covariance matrix from the ensemble averages obtained from the sigma points:

$$\boldsymbol{K}_k = \boldsymbol{P}_{xz,k|k-1} \boldsymbol{P}_{z,k|k-1}^{-1}, \tag{18}$$

$$\boldsymbol{\mu}_k = \boldsymbol{\mu}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{z}_{k|k-1}), \tag{19}$$

$$\boldsymbol{P}_k = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \boldsymbol{P}_{z,k|k-1} \boldsymbol{K}_k^T. \tag{20}$$

In the above UKF, steps 1 to 3 (from eq. (5) to eq. (10)) are known as the time update and the remaining steps are referred to as the measurement update. While the UKF assumes the state noise $\boldsymbol{w}_k$ and measurement noise $\boldsymbol{v}_k$ to be Gaussian, we consider a new BUKF-GM, which allows the noises to be non-Gaussian and they can be approximated using GMs.

2.2 Bayesian UKF with Gaussian Mixture (BUKF-GM)

In the proposed BUKF-GM, the pdfs of the state noise $\boldsymbol{w}_k$ and observation noise $\boldsymbol{v}_k$ are characterized by finite GMs:

$$p(\boldsymbol{w}_k) = \sum_{i=1}^{I} \alpha_{ik} N(\boldsymbol{w}_k, \overline{\boldsymbol{w}}_{ik}, \boldsymbol{Q}_{ik}), \tag{21}$$

$$p(\boldsymbol{v}_k) = \sum_{j=1}^{J} \beta_{jk} N(\boldsymbol{v}_k, \bar{\boldsymbol{v}}_{jk}, \boldsymbol{R}_{jk}), \quad (22)$$

where $I(J)$ is the number of components of the corresponding noise, $\alpha_{ik}(\beta_{jk})$ is the probability of the corresponding component with $\sum_{i=1}^{I}\alpha_{ik}=1$ ($\sum_{j=1}^{J}\beta_{jk}=1$). Now suppose further that the previous posteriori pdf of the state is modeled by the following GM with G components:

$$p(\boldsymbol{x}_{k-1}|\boldsymbol{z}_{k-1}) = \sum_{g=1}^{G} \gamma_{g(k-1)} N(\boldsymbol{x}_{k-1}, \boldsymbol{\mu}_{g(k-1)}, \boldsymbol{P}_{g(k-1)}), \quad (23)$$

where $\gamma_{g(k-1)}$ is the weight of the corresponding component. Then the predictive a priori density can be obtained as:

$$\begin{aligned} p(\boldsymbol{x}_k|\boldsymbol{z}_{k-1}) &= \int p(\boldsymbol{x}_{k-1}|\boldsymbol{z}_{k-1}) p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) d\boldsymbol{x}_{k-1} \\ &= \sum_{g'=1}^{G'} \gamma_{g'(k|k-1)} N(\boldsymbol{x}_k, \boldsymbol{\mu}_{g'(k|k-1)}, \boldsymbol{P}_{g'(k|k-1)}), \end{aligned} \quad (24)$$

where $G' = GI$, $\gamma_{g'(k|k-1)} = \gamma_{g(k-1)}\alpha_{ik} \Big/ \sum_{g=1}^{G}\sum_{i=1}^{I}\gamma_{g(k-1)}\alpha_{ik}$, and $\boldsymbol{\mu}_{g'(k|k-1)}$ and $\boldsymbol{P}_{g'(k|k-1)}$ are respectively the mean and covariance of corresponding Gaussian component, which can be predicted using the time update of UKF from eq. (5) to eq. (10).

Furthermore, given the current observation at time instant k, the corrected a posteriori density can be derived as follows:

$$\begin{aligned} p(\boldsymbol{x}_k|\boldsymbol{z}_k) &= c_k\, p(\boldsymbol{z}_k|\boldsymbol{x}_k) p(\boldsymbol{x}_k|\boldsymbol{z}_{k-1}) \\ &= c_k \sum_{j=1}^{J} \beta_{jk} N(\boldsymbol{z}_k, h_k(\boldsymbol{x}_k), \boldsymbol{R}_{jk}) \cdot \\ &\quad \sum_{g'=1}^{G'} \gamma_{g'(k|k-1)} N(\boldsymbol{x}_k, \boldsymbol{\mu}_{g'(k|k-1)}, \boldsymbol{P}_{g'(k|k-1)}) \\ &= \sum_{g''=1}^{G''} \gamma_{g''k} N(\boldsymbol{x}_k, \boldsymbol{\mu}_{g''k}, \boldsymbol{P}_{g''k}), \end{aligned} \quad (25)$$

where $c_k^{-1} = \int p(\boldsymbol{x}_k|\boldsymbol{z}_{k-1}) p(\boldsymbol{z}_k|\boldsymbol{x}_k) d\boldsymbol{x}_k$ represents a normalizing constant, $G'' = G'J = GIJ$ is the updated component number, $\gamma_{g''k} = \gamma_{g'(k|k-1)}\beta_{jk}\, p_j(\boldsymbol{z}_k|\boldsymbol{x}_k) \Big/ \sum_{g'=1}^{G'}\sum_{j=1}^{J}\gamma_{g'(k|k-1)}\beta_{jk}\, p_j(\boldsymbol{z}_k|\boldsymbol{x}_k)$, and $\boldsymbol{\mu}_{g''k}$ and $\boldsymbol{P}_{g''k}$ are respectively the mean and covariance of the corresponding Gaussian component, which can be estimated using the measurement update of UKF from eq. (11) to eq. (20). It can be noticed that the number of mixture components will grow from G to G' in the predictive step and from G' to G'' in the subsequent measurement correction step. Hence, the number of mixture components and hence the complexity will grow exponentially with time, which complicates online applications. We now address this issue by means of an order reduction approach, which help to maintain the number of components over time.

2.3 BUKF with Simplified GM (BUKF-SGM)

Consider the following GM model with n components:

$$f(\boldsymbol{x}) = \sum_{j=1}^{n} \alpha_j \phi_j(\boldsymbol{x}), \quad (26)$$

where $\phi_j(\boldsymbol{x}) = N(\boldsymbol{x}-\boldsymbol{u}_j, \boldsymbol{H}_j)$ is the j-th components and α_j are the mixing coefficients such that $\sum_{j=1}^{n}\alpha_j = 1$. In the BUKF-SGM, our goal is to approximate $f(\boldsymbol{x})$ as a simplified mixture model with fewer components $g(\boldsymbol{x}) = \sum_{i=1}^{m} w_i g_i(\boldsymbol{x})$, where $g_i(\boldsymbol{x}) = N(\boldsymbol{x}-\boldsymbol{t}_i, \widetilde{\boldsymbol{H}}_i)$ with $m<n$. w_i, $\boldsymbol{t}_i$ and $\widetilde{\boldsymbol{H}}_i$ are respectively the weight, center and covariance matrix of the i-th component $g_i(\boldsymbol{x})$. w_i 's should be summed up to one.

TABLE I BUKF-SGM ALGORITHM

Initialization

$p(x_0) = \sum_{g=1}^{G} \gamma_{g0} N(\boldsymbol{x}_0, \boldsymbol{\mu}_{g0}, \boldsymbol{P}_{g0})$

for $k = 1,2,...$

BUKF-SGM Prediction:

for $i = 1:I$

for $g = 1:G$

$g' = g + (i-1)G$

Estimate $\boldsymbol{\mu}_{g'(k|k-1)}$ and $\boldsymbol{P}_{g'(k|k-1)}$ with $\boldsymbol{\mu}_{g(k-1)}, \boldsymbol{P}_{g(k-1)}$ and $\boldsymbol{Q}_{ik}$ (UKF Time Update)

$\gamma^{*}_{g'(k|k-1)} = \gamma_{g(k-1)}\alpha_{ik}$

end

end

$\gamma_{g'(k|k-1)} = \gamma^{*}_{g'(k|k-1)} \Big/ (\sum_{g=1}^{G}\sum_{i=1}^{I}\gamma_{g(k-1)}\alpha_{ik})$

BUKF-SGM Correction:

for $j = 1:J$

for $g' = 1:GI$

$g'' = g' + (j-1)GI$

Estimate $\boldsymbol{\mu}_{g''k}$ and $\boldsymbol{P}_{g''k}$ with $\boldsymbol{\mu}_{g'(k|k-1)}$, $\boldsymbol{P}_{g'(k|k-1)}$ and $\boldsymbol{R}_{jk}$ (UKF Measurement Update)

$\gamma^{*}_{g''k} = \gamma_{g'(k|k-1)}\beta_{jk}\, p_j(\boldsymbol{z}_k|\boldsymbol{x}_k)$

end

end

$\gamma_{g''k} = \gamma^{*}_{g''k} \Big/ (\sum_{g'=1}^{G'}\sum_{j=1}^{J}\gamma_{g'(k|k-1)}\beta_{jk}\, p_j(\boldsymbol{z}_k|\boldsymbol{x}_k))$

BUKF-SGM Order Reduction:

Compute:

$\{\gamma_{gk}\}_{g=1}^{G}$, $\{\boldsymbol{\mu}_{gk}\}_{g=1}^{G}$ and $\{\boldsymbol{P}_{gk}\}_{g=1}^{G}$

end

Given a distance metric $D(f(\boldsymbol{x}), g(\boldsymbol{x}))$ between functions $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$, the error of approximating $f(\boldsymbol{x})$ with $g(\boldsymbol{x})$ is $D(f(\boldsymbol{x}), g(\boldsymbol{x})) = \left(\int (f(\boldsymbol{x}) - g(\boldsymbol{x}))^2 d\boldsymbol{x}\right)^{0.5}$. Different from the conventional re-sampling and clustering techniques, we adopted the two-step GM simplification procedure developed in [11]. More precisely, at the k-th iteration, the component mixture is partitioned into m groups $\{S_1^{(k)}, S_2^{(k)}, ..., S_m^{(k)}\}$.

Step 1 (Mean update): The representative component $C_i^{(k)}(\boldsymbol{x})$ for $S_i^{(k)}$ that minimizes the local quantization error is $C_i^{(k)}(\boldsymbol{x}) = \arg\min_{C(\boldsymbol{x})} \sum_{j\in S_i^{(k)}} \alpha_j \int (C(\boldsymbol{x}) - \phi_j(\boldsymbol{x}))^2 d\boldsymbol{x}$. Interested readers are referred to [11] and references therein for solving this problem using coordinate descent.

Step 2 (Clustering): Given $C^{(k)} = \{C_i^{(k)}(\boldsymbol{x})\}_{i=1}^{m}$, one re-assign $\phi_j(\boldsymbol{x})$ to the nearest $C_i^{(k)}(\boldsymbol{x})$ based on the distortion measure $D_{i,j} = D(C_i^{(k)}(\boldsymbol{x}), \phi_j(\boldsymbol{x}))$, and then update $S_i^{(k)}$.

The above process is repeated until either 1) the change in total distortion or $C^{(k)} = \{C_i^{(k)}(\boldsymbol{x})\}_{i=1}^{m}$ is less than a certain threshold, or 2) a maximum number of iteration is reached.

Since the unscented transformation helps us to approximate the nonlinear density for each component after

the nonlinear measurement and state transformation, a set of Gaussian components are obtained at the end of each iteration, which can be simplified using the GM simplification procedure with fewer components so that the number of components after each iteration can be maintained at a constant level. This leads to a great reduction in complexity in comparing with conventional re-sampling approaches. For instance, if the component number and the dimension of the state is respectively n and d, the complexity of conventional approach such as the KLD-based model order reduction method in [7] is $O(1000n^2d^2)$, while the complexity of the function approximation-based method is $O(Td^3n(L+m)$, where T and L are the numbers of iterations which are typically small [11]. Table I summarizes the proposed BUKF-SGM algorithm.

III. Simulation Results

We consider the following nonlinear and non-Gaussian time series estimation problem [7], which is commonly used in the literature. The process model is given by:

$$x_k = 1 + \sin(\omega\pi(k-1)) + \phi_1 x_{k-1} + w_k, \quad (27)$$

where w_k has a Gamma pdf, $p(w_k) = Ga(3,0.5)$, and $\omega = 0.04$ and $\phi_1 = 0.5$ are scalar parameters. Data samples are assumed to be taken from k=1 to 60. The underlying state distribution can be heavy tailed and asymmetric due to the Gamma distributed process noise and the nonlinearity of the state function. The non-stationary observation model is defined as:

$$z_k = \begin{cases} \phi_2 x_k^2 + v_k, & k \le 30, \\ \phi_3 x_k - 2 + v_k, & k > 30, \end{cases} \quad (28)$$

where $\phi_2 = 0.2$ and $\phi_3 = 0.5$. The measurement noise is a zero-mean Gaussian distribution, i.e., $v_k \sim N(v_k, 0, 10^{-5})$.

We compare the performance of the proposed BUKF-SGM algorithm with the standard PF, GSPF [5], SPPF [12] and GMSPPF [7] in terms of the following root mean squared error (RMSE) measure, which is defined as:

$$RMSE_k = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(x_k(m) - \hat{x}_k(m))^2}, \quad (29)$$

where $x_k(m)$ and $\hat{x}_k(m)$ denote respectively the true and filtered state at the m-th Monte Carlo (MC) run. The settings are: 1) For the standard PF, the number of particles is 2000 and re-sampling is performed with a threshold value of 0.001. 2) For GSPF, the initial state pdf is modeled by five GMs with equal weights, i.e., $p(x_0) = \sum_{g=1}^{5} 0.2N(\boldsymbol{x}_0, 0, 10)$ and the Gamma pdf of the process noise is approximated by two GMs using the EM algorithm. The number of particles for GSPF is the same as the standard PF. 3) For SPPF, GMSPPF and BUKF-SMG, the initial state pdf is identical to that in the GSPF while the scaling parameter κ for sigma point generation is set to the recommended value with $\kappa = 3 - L = 2$. 4) The number of particles for GSPF, SPPF and GMSPPF is the same as the standard PF.

Experiments were conducted on a computer with Intel(R) Core(TM) i7-2600K CPU and 8192MB RAM. The RMSE averaged over 100 Monte Carlo trials for different algorithms is compared in Fig. 1. The RMSE values in Fig. 1 are further averaged over time and are shown in Table II. The results demonstrate that the proposed BUKF-SGM algorithm is able to obtain better estimation accuracy in comparing with the other algorithms. The performance of PF is similar with [7]. The reason for its bad performance is mainly due to the highly peaked likelihood function of the observations (arising from the small observation noise variance) fused with the spurious jumps in the state introduced by the heavy tailed process noise, which causes degeneracy and sampling impoverishment for the PF. GSPF, SPPF and GMSPPF can handle the case properly. However their accuracy mainly relies on the number of sampling particles and as we mentioned before the computational complexity will grow rapidly when the dimension of the states or the number of particles becomes larger. Compared with these PF-based algorithms, the overall performance of the proposed BUKF-SGM is better and its computational time averaged over 100 MC trials is also relatively lower as shown in table II.

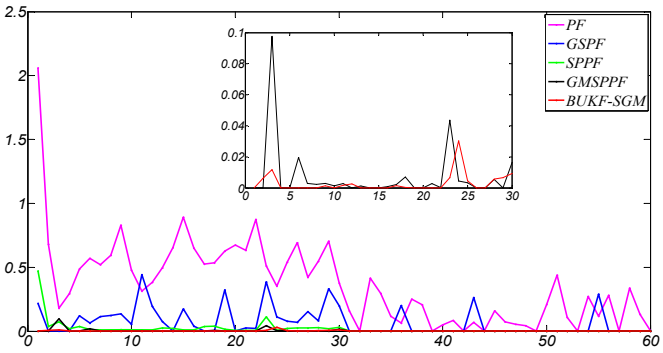

Fig. 1. RMSE of the estimated states for different algorithms are compared. An enlarged figure for the comparison between GMSPPF and BUKF-SGM from time instant 0 to 30 is also plotted.

TABLE II AVERAGE RMSE/TIME COMPARISON

	PF	GSPF [5]	SPPF [12]	GMSPPF [7]	**BUKF-SGM**
Average RMSE	0.2191	0.0465	0.0286	0.0063	**0.0016**
Average Time (s)	2.3506	1.6391	56.5715	1.1590	**0.8269**

IV. Conclusion

A novel BUKF algorithm with simplified Gaussian mixtures (BUKF-SGM) has been presented for dynamic state space estimation in nonlinear and non-Gaussian systems. Simulations show that the proposed BUKF-SGM algorithm can achieve better performance than the PF-based algorithms, which provides an attractive alternative to conventional nonlinear state estimation algorithms.

Reference

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Eng.*, vol. 82D, pp.35-45, Mar. 1960.

[2] J. Uhlmann *et al*, "A new method for the non linear transformation of means and covariances in filters and estimations," *IEEE Trans. on Automatic Contol*, vol. 45, 2000.

[3] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. of Amer. Stat. Assoc.*,vol. 93, no. 443, pp. 1032-1044. 1998.

[4] H. Moradkhani *et al*, "Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter," *Water Resources Research*, vol. 41, no.5, 2005.

[5] J. H. Kotecha and P. M. Djurić, "Gaussian sum particle filtering for dynamic state space models," in *Proc. IEEE Int. Conf. Acoust. Speech and Signal Process.,* 2001, pp. 3465-3468.

[6] I. Bilik and J. Tabrikian, "MMSE-based filtering in presence of non-Gaussian system and measurement noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, pp. 1153-1170, Jul. 2010.

[7] R. van der Merwe and E. Wan, "Gaussian mixture sigma-point particle filters for sequential probabilistic inference in dynamic state-space models," *in Proc. IEEE Int. Acoust., Speech, and Signal Process., 2003.* pp. VI-701.

[8] J. J. Verbeek *et al*, "Efficient greedy learning of Gaussian mixture models," *Neural Computation*, vol. 15, pp. 469-485, 2003.

[9] S. C. Chan *et al*, "Bayesian Kalman filtering, regularization and compressed sampling," *in Proc. IEEE Int. 54th Int. Midwest Symp. Circuits Syst.*, 2001, pp. 1-4.

[10] S. Zhang *et al*, "New Object Detection, Tracking, and Recognition Approaches for Video Surveillance Over Camera Network," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2679-2691, 2015.

[11] K. Zhang and J. T. Kwok, "Simplifying mixture models through function approximation," *IEEE Trans. Nerual Networks*, vol. 21, no. 4, pp. 644-658, Apr. 2010.

[12] R. Van Der Merwe *et al*, "The unscented particle filter," *Neural Inform. Process. Syst.*, vol. 2000, pp. 584-590, 2000.