

Received May 30, 2020, accepted June 19, 2020, date of publication June 25, 2020, date of current version July 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3004804

# Adaptive Fading Bayesian Unscented Kalman Filter and Smoother for State Estimation of Unmanned Aircraft Systems

ZHONG LIU<sup>1,2</sup> AND SHING-CHOW CHAN<sup>2</sup>, (Member, IEEE)

<sup>1</sup>Department of Biomedical Engineering, Shenzhen University, Shenzhen 518060, China

<sup>2</sup>Department of Electrical and Electronics Engineering, The University of Hong Kong, Hong Kong

Corresponding author: Zhong Liu (liuzhong@szu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61901282.

**ABSTRACT** This paper proposes an adaptive fading Bayesian unscented Kalman filter (AF-BUKF) and explores its application for state estimation of unmanned aircraft systems (UASs). In the AF-BUKF, 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 prohibitive computational complexity caused by the exponential growth of mixture components, a Gaussian mixture simplification algorithm is employed. Moreover, the AF-BUKF algorithm employs a novel adaptive fading strategy to recursively update the Gaussian components, so that the adverse effect of inexact knowledge of the state and measurement noise covariance can be mitigated. An AF-BUK Smoother (AF-BUKS) is also proposed by extending the AF-BUKF algorithm using the concept of optimal Bayesian smoothing and the Rauch-Tung-Striebel Smoother to improve estimation accuracy. Experimental results on simulated and real UAS data show that the proposed AF-BUKF/S algorithms can achieve better performance compared with the conventional methods. Thus, they can serve as attractive alternative approaches for nonlinear state estimation of UASs and other problems.

**INDEX TERMS** Bayesian smoothing, nonlinear and non-Gaussian system, Gaussian mixture, unmanned aircraft systems, unscented Kalman filter.

## I. INTRODUCTION

Small unmanned aircraft systems (UASs) [1]–[4] have gained growing popularity owing to their convenience for mobile missions such as search and rescue [5], monitoring [6], path planning [7], [8], and target tracking [9]–[11]. The control of UASs requires the prior knowledge of the corresponding state parameters which may include the positions, velocity, orientation and orientation rate of the aircraft. Though some of the parameters can be measured using specialized sensors such as Global Positioning System (GPS) receivers and Inertial Measurement Units (IMUs) [12], the signals obtained from such sensors may be contaminated by noise/interference from ambient environments and the bias inherited in the sensors themselves. Therefore, it is important to develop appropriate onboard algorithms for estimating the states of UASs from noisy measurements.

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy<sup>1</sup>.

Kalman filter (KF) and its various extensions have been widely utilized in state estimation since the dynamic model of the system states can be incorporated through the state space model for tracking together with their measurements. Consequently, the estimation performance can be improved by taking advantage of the additional prior information provided by the state dynamic. Moreover, uncertainties in the system model and measurements can be conveniently incorporated via the additive noises in the model.

In the traditional KF [13], the system model, which includes the state equation and measurement equation, is assumed to be linear. Moreover, the modeling error or noise in the state equation and measurement noise in the measurement equation are assumed to be zero mean Gaussian distributed with known covariance matrices. Under these conditions, the KF is the optimal estimator with minimum mean squares errors. However, for more complicated systems, both the state and measurement equations are nonlinear functions. Furthermore, the uncertainties encapsulated in form of modeling and

measurement noises may be non-Gaussian. Thus, in this case, there are generally no analytic expressions for the *a posteriori* densities, and various extensions of the KF have been developed towards this end. For general nonlinear systems with non-Gaussian densities, the state and noise densities can be approximated by various methods such as histogram or Gaussian mixtures (GMs). The nonlinearity in the state and measurement equations are usually tackled through linearization as in the extended Kalman filter (EKF) [11], [14], [15] or by approximating the state density after the transformation using Gaussian distribution through unscented transformation as in the unscented KF (UKF) [16]–[18]. Other approximations include the cubature KF [19] and Gauss-Hermite KF [20].

The Monte Carlo (MC) statistical approach represents the density as a set of particles and it gives rise to the particle filter (PF) technique [21], [22], which approximates the non-Gaussian densities with a set of weighted particles. However, the computational complexity of the MC statistical method will grow exponentially with the dimension of the states and particles. To address this problem, many GM-based PFs [22]–[24] have been proposed but the number of components in the GM will grow rapidly with the number of iterations. To avoid this growing complexity, re-sampling and clustering techniques have been used to simplify the resultant density at each iteration. These developments paved the way for filtering algorithms to solve more challenging problems such as the simultaneous localization and mapping problem [16]. Furthermore, various algorithms have been proposed for KF filtering with heavy-tailed non-Gaussian noises, such as Student's  $t$  UKF (STUKF) [25], Student's  $t$  cubature KF (STCKF) [26], robust KF [27] and other techniques described therein.

In our previous work [28], a Bayesian Kalman filter (BKF) for linear systems with non-Gaussian noise was proposed. It approximates the non-Gaussian density as a Gaussian mixture and uses a novel Gaussian mixture simplification algorithm to maintain the number of components in the mixture at a reasonable level. This helps to keep the complexity of the resultant BKF at a reasonably low level. The effectiveness of the BKF algorithm has been demonstrated in video object tracking and other applications [29], [30]. The usefulness of the non-Gaussian state and measurement noises is their ability to model sudden change in system dynamics and possible measurement outliers.

In applications such as small unmanned aircraft systems (UASs) [1]–[4], the system models can be highly nonlinear [3], [14] and may subject to various disturbances. Therefore, it would be desirable if the above linear BKF with simplified Gaussian mixtures can be extended to these nonlinear settings. Another major practical issue in KFs is the selection of the covariance matrices of the modeling and measurements noises. This is also true for the Gaussian mixtures in the BKF where the covariance matrices of the GMs need to be determined *a priori*.

To address these important challenges, this paper aims to extend the BKF concept in the following perspectives:

1) *State estimation of nonlinear systems with non-Gaussian distributed noises*: as mentioned earlier, the BKF algorithm offers promising performance for state estimation of linear non-Gaussian systems. It is desirable to extend these nice properties to solve nonlinear system identification problems.

2) *Adaptive estimation of GM parameters under modeling and time-varying model errors*: the BKF algorithm offers a framework for state estimation using a simplified GM model. It works well when the state and measurement noises are perfectly known *a priori*. However, in practical situations, such quantities may not be known exactly or even be time-varying. Moreover, it is usually tedious to select such parameters by trial and error. In time-varying scenarios, the performance of the BKF with these preset values may also be compromised. Therefore, it is necessary to devise an automatic method for updating adaptively the parameters of the GMs so that such modeling or time-varying errors can be mitigated.

To this end, we propose a novel adaptive fading Bayesian unscented Kalman filter/smoother with simplified Gaussian mixtures (AF-BUKF/S) for state estimation in nonlinear and non-Gaussian systems. Among different nonlinear filtering methods, we have chosen the unscented transform<sup>1</sup> as our starting point due to its outstanding 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, which directly simplifies the GMs by minimizing an upper bound of the approximation error between the original and the simplified models. The resultant filter is similar to the Gaussian sum unscented Kalman filter (GSUKF) [31], [32], except that the noise covariances are made adaptive by utilizing the adaptive fading approach and the BKF-SGM method in [28]–[30] is used for mixture reduction.<sup>2</sup> The adaptive fading approach<sup>3</sup> estimates a correction factor for the nominal noise covariance matrices so as to account for possible mismatch in these parameters, as they may not be exactly known or even be time varying. Furthermore, the correction parameters are tracked using BKF-SGM using a random walk model to further reduce its variance and improve its tracking speed. This greatly extends the applicability of the BUKF-SGM and yields improved estimation performance. Finally, the AF-BUKF method is further extended to an AF-BUKS algorithm by employing the optimal Bayesian smoothing formula and the

<sup>1</sup>The proposed approach is also applicable to the cubature KF [19] and Gauss-Hermite KF [20]. Here, we shall present the details on UKF.

<sup>2</sup>The usefulness of this concept was briefly introduced using simulation in our preliminary work [33]. Here we further improve its adaptability by using the adaptive fading approach and propose its smoother counterpart. Its application to the state estimation of UASs is demonstrated by detailed simulations.

<sup>3</sup>Due to page limitations, previous works on adaptive fading estimation are reviewed in the supplementary materials where we show that the proposed adaptive fading approach can offer better performance than the other state-of-the-art algorithms in [34]–[37].

Rauch-Tung-Striebel smoother concept to improve estimation accuracy by using future measurements. The AF-BUKS algorithm is useful when finite delays can be tolerated by utilizing certain amount of future data for improving the current estimation. Experimental results and comparisons on simulated UASs datasets and real flight data show that the proposed AF-BUKF/S algorithm can offer significantly better performance for state estimation of UASs than the conventional approaches.

In summary, our contributions are:

- We extended the BKF-SGM to state estimation in nonlinear and non-Gaussian systems using the UKF and the Gaussian mixture representation of the excitation and measurement noises. It utilizes the Gaussian mixture simplification algorithm to avoid the exponential growth in complexity over time and the resultant BUKF-SGM algorithm gives better estimation accuracy than the particle filter (PF) and other GM simplification methods including the pruning and Kullback-Leibler-based merging (KLM) approaches.
- A novel adaptive fading method to update the scale parameter of the Gaussian component, which is proposed for the BUKF-SGM framework to give an AF-BUKF algorithm for better estimation of the GM parameters to account for modeling and time-varying model errors.
- To further improve the estimation accuracy, the AF-BUKF algorithm is extended by means of an optimal Bayesian smoothing concept, yielding an AF-BUKS algorithm which can be used in offline applications or online applications where certain delay can be tolerated. Moreover, it yields better performance than the AF-BUKF. Finally, the application of the proposed algorithms to UASs is studied through extensive simulations on real and synthetic data under both non-Gaussian excitation and measurement noises.

The rest of the paper is organized as follows: we briefly review the background and the conventional UKF method in Section II. The proposed AF-BUKF/S algorithms are described in Section III. In Section IV, the performance of the proposed methods for state estimation of UASs are evaluated and compared with the conventional algorithms including EKF and UKF. Section V concludes the paper.

## II. BACKGROUND AND UKF

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

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}) + \boldsymbol{\omega}_k, \quad (1a)$$

$$\mathbf{z}_k = h_k(\mathbf{x}_k) + \mathbf{v}_k, \quad (1b)$$

where  $\mathbf{x}_k \in \mathbb{R}^L$  and  $\mathbf{z}_k \in \mathbb{R}^m$  denote respectively the state and measurement vectors at time instant  $k$ ,  $f_k(\cdot)$  and  $h_k(\cdot)$  are respectively known nonlinear state and measurement functions, and  $\boldsymbol{\omega}_k$  and  $\mathbf{v}_k$  represent respectively the state and measurement noises, which are independent of each other. We first introduce in this section the conventional

UKF where  $\boldsymbol{\omega}_k$  and  $\mathbf{v}_k$  are Gaussian distributed, whereas in Section III, we shall describe the proposed Bayesian UKF with Gaussian Mixture (BUKF-GM) by allowing the noises to be non-Gaussian distributed. The adaptive fading BUKF (AF-BUKF) for adaptively updating the covariance matrices of the GM will also be described.

The UKF algorithm belongs to a general class of nonlinear KF filter called Gaussian filtering [36], which assumes that the probability density functions (pdfs) of the state noise  $\boldsymbol{\omega}_k$  and measurement noise  $\mathbf{v}_k$  to be Gaussian distributed, i.e.,

$$p(\boldsymbol{\omega}_k) = \mathcal{N}(\boldsymbol{\omega}_k; \bar{\boldsymbol{\omega}}_k, \mathbf{Q}_k), \quad (2a)$$

$$p(\mathbf{v}_k) = \mathcal{N}(\mathbf{v}_k; \bar{\mathbf{v}}_k, \mathbf{R}_k), \quad (2b)$$

where  $\mathcal{N}(\mathbf{u}_k; \bar{\mathbf{u}}_k, \mathbf{C}_k)$  denotes a Gaussian distribution with mean  $\bar{\mathbf{u}}_k$  and covariance  $\mathbf{C}_k$ . Usually, the noises are of zero mean and hence  $\bar{\boldsymbol{\omega}}_k = 0$  and  $\bar{\mathbf{v}}_k = 0$ . Consequently, using (1a), one gets the predicted state mean and covariance as:

$$\begin{aligned} \boldsymbol{\mu}_{x,k|k-1} &= \mathbb{E}[\mathbf{x}_k] = \mathbb{E}[f_k(\mathbf{x}_{k-1})] + \mathbb{E}[\boldsymbol{\omega}_k] \\ &= \int f_k(\mathbf{x}_{k-1}) \mathcal{N}(\mathbf{x}_{k-1}; \boldsymbol{\mu}_{x,k-1}, \mathbf{P}_{x,k-1}) d\mathbf{x}_{k-1} + \bar{\boldsymbol{\omega}}_k, \end{aligned} \quad (3a)$$

$$\begin{aligned} \mathbf{P}_{x,k|k-1} &= \text{cov}[\mathbf{x}_k] = \text{cov}[f_k(\mathbf{x}_{k-1})] + \text{cov}[\boldsymbol{\omega}_k] \\ &= \int (f_k(\mathbf{x}_{k-1}) - \mathbb{E}[f_k(\mathbf{x}_{k-1})]) \\ &\quad \times (f_k(\mathbf{x}_{k-1}) - \mathbb{E}[f_k(\mathbf{x}_{k-1})])^T \\ &\quad \times \mathcal{N}(\mathbf{x}_{k-1}; \boldsymbol{\mu}_{x,k-1}, \mathbf{P}_{x,k-1}) d\mathbf{x}_{k-1} + \mathbf{Q}_k \\ &= \int (f_k(\mathbf{x}_{k-1}) + \bar{\boldsymbol{\omega}}_k - \boldsymbol{\mu}_{x,k|k-1}) \\ &\quad \times (f_k(\mathbf{x}_{k-1}) + \bar{\boldsymbol{\omega}}_k - \boldsymbol{\mu}_{x,k|k-1})^T \\ &\quad \times \mathcal{N}(\mathbf{x}_{k-1}; \boldsymbol{\mu}_{x,k-1}, \mathbf{P}_{x,k-1}) d\mathbf{x}_{k-1} + \mathbf{Q}_k. \end{aligned} \quad (3b)$$

Similarly, according to (1b), the predicted measurement mean, covariance as well as the cross-covariance of the state and measurement can be obtained as follows:

$$\begin{aligned} \boldsymbol{\mu}_{z,k|k-1} &= \mathbb{E}[\mathbf{z}_k] = \mathbb{E}[h_k(\mathbf{x}_k)] + \mathbb{E}[\mathbf{v}_k] \\ &= \int h_k(\mathbf{x}_k) \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{x,k|k-1}, \mathbf{P}_{x,k|k-1}) d\mathbf{x}_k + \bar{\mathbf{v}}_k, \end{aligned} \quad (4a)$$

$$\begin{aligned} \mathbf{P}_{z,k|k-1} &= \text{cov}[\mathbf{z}_k] = \text{cov}[h_k(\mathbf{x}_k)] + \text{cov}[\mathbf{v}_k] \\ &= \int (h_k(\mathbf{x}_k) + \bar{\mathbf{v}}_k - \boldsymbol{\mu}_{z,k|k-1}) \\ &\quad \times (h_k(\mathbf{x}_k) + \bar{\mathbf{v}}_k - \boldsymbol{\mu}_{z,k|k-1})^T \\ &\quad \times \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{x,k|k-1}, \mathbf{P}_{x,k|k-1}) d\mathbf{x}_k + \mathbf{R}_k, \end{aligned} \quad (4b)$$

$$\begin{aligned} \mathbf{P}_{xz,k|k-1} &= \int (\mathbf{x}_k - \boldsymbol{\mu}_{x,k|k-1}) \\ &\quad \times (h_k(\mathbf{x}_k) + \bar{\mathbf{v}}_k - \boldsymbol{\mu}_{z,k|k-1})^T \\ &\quad \times \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{x,k|k-1}, \mathbf{P}_{x,k|k-1}) d\mathbf{x}_k. \end{aligned} \quad (4c)$$

Finally, one can update the state mean and error covariance as:

$$\mathbf{K}_k = \mathbf{P}_{xz,k|k-1} \mathbf{P}_{z,k|k-1}^{-1}, \quad (5a)$$

$$\boldsymbol{\mu}_k = \boldsymbol{\mu}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \boldsymbol{\mu}_{z,k|k-1}), \quad (5b)$$

$$\mathbf{P}_{x,k} = \mathbf{P}_{x,k|k-1} - \mathbf{K}_k \mathbf{P}_{z,k|k-1} \mathbf{K}_k^T. \quad (5c)$$

In the conventional EKF,  $f_k(\mathbf{x}_k)$  and  $h_k(\mathbf{x}_k)$  are locally linearized at  $\mathbf{x}_k$  to obtain a locally linear state space system. Its limitation is that the error resulting from the linearization will usually increase with the step taken. Hence, one would expect that when a large step is taken at region with strong nonlinearity, the error in tracking the states will be large. 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. Other methods to approximate the integrals in (3b) and (4b) include cubature KF [19], particle filters [21] and Gauss-Hermite KF [20]. It was shown in [38] that the performances of the cubature KF, UKF and Gauss-Hermite KF are similar if the covariance is estimated by the variational Bayes approach. Therefore, we shall focus on the UKF in this paper.

Given an  $L$ -dimensional state vector  $\mathbf{x}_k$  with mean  $\boldsymbol{\mu}_k$  and covariance  $\mathbf{P}_k$ , the UKF algorithm [16], [18] generates two samples, so-called sigma points, along each dimension and evaluates the outputs of the corresponding nonlinearity. By doing so, one can estimate the Gaussian approximation of the state density by estimating empirically the corresponding mean and covariance from the samples at the sigma points. For instance, given the Gaussian approximated state density at time  $k-1$ ,  $p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$ , where  $\mathbf{Z}_{k-1} = \{\mathbf{z}_1, \dots, \mathbf{z}_{k-1}\}$  denotes the measurements up to the  $(k-1)$ -th time instant, one can update the predicted density  $p(\mathbf{x}_k|\mathbf{x}_{k-1}) = p_{f,w_k}(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$  from the dynamic equation through sampling, where  $p_{f,w_k}(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$  denotes the density after the nonlinear transformation of  $p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$  with  $f(\cdot)$  followed by the addition of  $\boldsymbol{\omega}_k \propto \mathcal{N}(\boldsymbol{\omega}_k; \bar{\boldsymbol{\omega}}_k, \mathbf{Q}_k)$ . To this end, samples of  $p_{f,w_k}(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$  at the sigma points are computed and thus  $p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})$  is approximated by a Gaussian distribution empirically. This gives the time update in Algorithm 1. The sampling technique can also be used to update the state density  $p(\mathbf{x}_k|\mathbf{Z}_k)$  from  $p(\mathbf{x}_k|\mathbf{x}_{k-1})$  via the measurement equation in (1b), where  $\mathbf{Z}_k = \{\mathbf{z}_1, \dots, \mathbf{z}_k\}$  denotes the measurements up to time instant  $k$ . This yields the measurement update in Algorithm 1.

Next, we shall propose a new Bayesian UKF with GM (BKF-GM) algorithm, which allows the noises to be GMs.

### III. THE PROPOSED BAYESIAN UKF/S

We first briefly outline the notation of the BUKF in Section A below, which is also similar to the GSUKF [31], [32]. As mentioned in the previous section, our development is also applicable to the cubature KF [19] and Gauss-Hermite KF [20]. The proposed mixture simplification algorithm will be introduced in Section III-B. The adaptive fading

### Algorithm 1 The UKF Algorithm

- 1: **Time Update**
- 2: Generate sigma points  $\{\boldsymbol{\chi}_{l,k-1}\}$ ,  $l = 1, \dots, L$ , with weights  $\{w_l\}$  from  $p(\mathbf{x}_{k-1})$ :
- 3:  $\boldsymbol{\chi}_{0,k-1} = \boldsymbol{\mu}_{x,k-1}$ ,  $w_0 = \frac{\kappa}{L+\kappa}$ ,
- 4:  $\boldsymbol{\chi}_{l,k-1} = \boldsymbol{\mu}_{x,k-1} + (\sqrt{(L+\kappa)\mathbf{P}_{x,k-1}})_l$ ,  $w_l = \frac{1}{2(L+\kappa)}$ ,
- 5:  $\boldsymbol{\chi}_{L+l,k-1} = \boldsymbol{\mu}_{x,k-1} - (\sqrt{(L+\kappa)\mathbf{P}_{x,k-1}})_l$ ,  $w_{L+l} = w_l$ ,
- 6: where  $\kappa$  is the scaling parameter,  $(\sqrt{\mathbf{A}})_l$  denotes the  $l$ -th column of the matrix square root of  $\mathbf{A}$ .
- 7: Propagate each sigma point:
- 8:  $\boldsymbol{\chi}_{l,k|k-1} = f_k(\boldsymbol{\chi}_{l,k-1})$ ,  $l = 0, \dots, 2L$ .
- 9: Approximate the mean and covariance for  $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ :
- 10:  $\boldsymbol{\mu}_{k|k-1} = \sum_{l=0}^{2L} w_l \boldsymbol{\chi}_{l,k|k-1}$ ,
- 11:  $\mathbf{P}_{x,k|k-1} = \sum_{l=0}^{2L} w_l (\boldsymbol{\chi}_{l,k|k-1} - \boldsymbol{\mu}_{x,k|k-1})$
- 12:  $\quad \times (\boldsymbol{\chi}_{l,k|k-1} - \boldsymbol{\mu}_{x,k|k-1})^T + \mathbf{Q}_k$ .
- 13: **Measurement Update**
- 14: Generate new sigma points with  $\boldsymbol{\mu}_{k|k-1}$  and  $\mathbf{P}_{k,k|k-1}$ :
- 15:  $\boldsymbol{\chi}_{0,k|k-1} = \boldsymbol{\mu}_{x,k|k-1}$ ,  $w_0 = \frac{\kappa}{L+\kappa}$ ,
- 16:  $\boldsymbol{\chi}_{l,k|k-1} = \boldsymbol{\mu}_{x,k|k-1} + (\sqrt{(L+\kappa)\mathbf{P}_{x,k|k-1}})_l$ ,  $w_l = \frac{1}{2(L+\kappa)}$ ,
- 17:  $\boldsymbol{\chi}_{L+l,k|k-1} = \boldsymbol{\mu}_{x,k|k-1} - (\sqrt{(L+\kappa)\mathbf{P}_{x,k|k-1}})_l$ ,  $w_{L+l} = w_l$ ,
- 18: where  $l = 1, \dots, L$ .
- 19: Estimate the measurement mean:
- 20:  $\mathbf{z}_{l,k|k-1} = h_k(\boldsymbol{\chi}_{l,k|k-1})$ ,  $l = 0, \dots, 2L$ ,
- 21:  $\boldsymbol{\mu}_{z,k|k-1} = \sum_{l=0}^{2L} w_l \mathbf{z}_{l,k|k-1}$ .
- 22: Estimate the measurement error covariance and cross-covariance:
- 23:  $\mathbf{P}_{zz,k|k-1} = \sum_{l=0}^{2L} w_l (\mathbf{z}_{l,k|k-1} - \boldsymbol{\mu}_{z,k|k-1})$
- 24:  $\quad \times (\mathbf{z}_{l,k|k-1} - \boldsymbol{\mu}_{z,k|k-1})^T$ ,
- 25:  $\mathbf{P}_{z,k|k-1} = \mathbf{P}_{zz,k|k-1} + \mathbf{R}_k$ ,
- 26:  $\mathbf{P}_{xz,k|k-1} = \sum_{l=0}^{2L} w_l (\boldsymbol{\chi}_{l,k|k-1} - \boldsymbol{\mu}_{z,k|k-1})$
- 27:  $\quad \times (\mathbf{z}_{l,k|k-1} - \boldsymbol{\mu}_{z,k|k-1})^T$ .
- 28: Estimate the Gaussian approximation of  $p(\mathbf{x}_k|\mathbf{Z}_k)$
- 29:  $\mathbf{e}_k = \mathbf{z}_k - \boldsymbol{\mu}_{z,k|k-1}$ ,
- 30:  $\mathbf{K}_k = \mathbf{P}_{xz,k|k-1} \mathbf{P}_{z,k|k-1}^{-1}$ ,
- 31:  $\boldsymbol{\mu}_{x,k} = \boldsymbol{\mu}_{x,k|k-1} + \mathbf{K}_k \mathbf{e}_k$ ,
- 32:  $\mathbf{P}_{x,k} = \mathbf{P}_{x,k|k-1} - \mathbf{K}_k \mathbf{P}_{z,k|k-1} \mathbf{K}_k^T$ .

BUKF will be introduced in Section III-C. The Adaptive Fading BUK Smoother (AF-BUKS) will be discussed in Section III-D.

#### A. BAYESIAN UKF (BUKF)

As mentioned earlier, the pdfs of the state noise  $\boldsymbol{\omega}_k$  and measurement noise  $\mathbf{v}_k$  in the proposed BUKF are characterized by a finite set of GMs as:

$$p(\boldsymbol{\omega}_k) = \sum_{i=1}^I \alpha_{i,k} \mathcal{N}(\boldsymbol{\omega}_k; \bar{\boldsymbol{\omega}}_{i,k}, \mathbf{Q}_{i,k}), \quad (6a)$$

$$p(\mathbf{v}_k) = \sum_{j=1}^J \beta_{j,k} \mathcal{N}(\mathbf{v}_k; \bar{\mathbf{v}}_{j,k}, \mathbf{R}_{j,k}), \quad (6b)$$

where  $I$  and  $J$  are respectively the numbers of components in the GMs of  $\boldsymbol{\omega}_k$  and  $\mathbf{v}_k$  with  $\alpha_{i,k}$  and  $\beta_{j,k}$  as the probabilities of the corresponding components in the GM, which satisfy  $\sum_{i=1}^I \alpha_{i,k} = 1$  and  $\sum_{j=1}^J \beta_{j,k} = 1$ . The GM in (6a) is useful to

model sudden change of states in the dynamical system such as directions in UASs. On the other hand, (6b) is useful to model measurement noise corrupted by impulsive noise and outliers. Since  $\omega_k$  and  $\mathbf{v}_k$  are now GMs, the pdf of the state is generally not Gaussian distributed. Hence, it is assumed that the *a posteriori* pdf of the state at the previous time instant is approximated by the following finite GM with  $G$  components:

$$p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1}) \approx \sum_{g=1}^G \gamma_{g,k-1} \mathcal{N}(\mathbf{x}_{k-1}; \boldsymbol{\mu}_{g,k-1}, \mathbf{P}_{g,k-1}), \quad (7)$$

where  $\gamma_{g,k-1}$ ,  $g = 1, \dots, G$ , denotes the weight of the  $g$ -th component at the  $k-1$  instant and  $\sum_{g=1}^G \gamma_{g,k-1} = 1$ . Consequently, the predictive priori density can be written as

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{Z}_{k-1}) &= \int p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1}) p(\mathbf{x}_k|\mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \\ &\approx \int \sum_{g=1}^G \gamma_{g,k-1} \mathcal{N}(\mathbf{x}_{k-1}; \boldsymbol{\mu}_{g,k-1}, \mathbf{P}_{g,k-1}) \\ &\quad \times \sum_{i=1}^I \alpha_{i,k} \mathcal{N}(\boldsymbol{\omega}_k; \mathbf{x}_k - f_k(\mathbf{x}_{k-1}) \\ &\quad - \bar{\boldsymbol{\omega}}_{i,k}, \mathbf{Q}_{i,k}) d\mathbf{x}_{k-1}, \end{aligned} \quad (8)$$

In (8), each combination of  $\mathcal{N}(\mathbf{x}_{k-1}; \boldsymbol{\mu}_{g,k-1}, \mathbf{P}_{g,k-1})$  and  $\mathcal{N}(\boldsymbol{\omega}_k; \mathbf{x}_k - f_k(\mathbf{x}_{k-1}) - \bar{\boldsymbol{\omega}}_{i,k}, \mathbf{Q}_{i,k})$  after the integration can be approximated as a Gaussian distribution  $\mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{g',k|k-1} + \bar{\boldsymbol{\omega}}_{g',k}, \mathbf{P}_{g',k|k-1})$ , where  $g' = 1, \dots, GI$ , is used to index all such combinations. The subscript  $k|k-1$  denotes the quantities in the predicted density of  $\mathbf{x}_k$  given measurements up to time instant  $k-1$ ,  $\mathbf{Z}_{k-1}$ . Consequently, for zero mean modeling noise, i.e.,  $\bar{\boldsymbol{\omega}}_{g',k} = 0$ , (8) can be approximately simplified to

$$\int \sum_{g'=1}^{G'} \gamma_{g',k|k-1} \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{g',k|k-1}, \mathbf{P}_{g',k|k-1}) d\mathbf{x}_{k-1}, \quad (9)$$

where  $G' = GI$  and  $\gamma_{g',k|k-1} = \gamma_{g,k-1} \alpha_{i,k}$ . The quantities  $\boldsymbol{\mu}_{g',k|k-1}$  and  $\mathbf{P}_{g',k|k-1}$ , for each  $g'$ -th component, can be predicted using the time update of the UKF in Algorithm 1.

Similarly, when the current observation at time instant  $k$  is available, the *a posteriori* density at time instant  $k$  can be updated as

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{Z}_k) &= c_k p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{Z}_{k-1}) \\ &= c_k \sum_{j=1}^J \beta_{j,k} \mathcal{N}(\mathbf{v}_k; \mathbf{z}_k - h_k(\mathbf{x}_k) - \bar{\mathbf{v}}_{j,k}, \mathbf{R}_{j,k}) \\ &\quad \times \sum_{g'=1}^{G'} \gamma_{g',k|k-1} \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{g',k|k-1}, \mathbf{P}_{g',k|k-1}) \\ &= \sum_{g''=1}^{G''} \gamma_{g'',k} \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{g'',k}, \mathbf{P}_{g'',k}), \end{aligned} \quad (10)$$

where  $c_k = (\int p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{Z}_{k-1}) d\mathbf{x}_k)^{-1}$  is a normalization constant,  $\gamma_{g'',k} = \frac{\gamma_{g',k|k-1} \beta_{j,k} p_j(\mathbf{z}_k|\boldsymbol{\mu}_{g'',k})}{\sum_{g''=1}^{G''} \sum_{j=1}^J \gamma_{g',k|k-1} \beta_{j,k} p_j(\mathbf{z}_k|\boldsymbol{\mu}_{g'',k})}$ , and  $G'' = G'I = GIJ$  is the number of components after the update.  $\boldsymbol{\mu}_{g'',k}$  and  $\mathbf{P}_{g'',k}$  are respectively the mean and covariance of the corresponding Gaussian component obtained with the prior component density  $\mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{g',k|k-1}, \mathbf{P}_{g',k|k-1})$  and

measurement  $\mathbf{z}_k$  and component noise covariance  $\mathbf{R}_{j,k}$ . This can be obtained with the measurement update of the UKF in Algorithm 1.

It can be seen that the number of mixture components will grow from  $G$  to  $G'$  in the prediction step and further from  $G'$  to  $G''$  in the subsequent measurement update step. Over time, the number of mixture components and hence the complexity will grow exponentially, which makes the computation prohibitive. Fortunately, this issue can be overcome by using the order reduction approach which we have proposed in [28] for the BKF. More precisely, the GM model in (10) is simplified to the form of (7) so that the number of mixture components is maintained after each pair of time and measurement update steps. For completeness, we shall briefly describe this order reduction process for GM simplification in the next subsection.

### B. BUKF WITH SIMPLIFIED GM (BUKF-SGM)

Consider the GM model in (10) with  $G''$  components:

$$p(\mathbf{x}) = \sum_{g''=1}^{G''} \gamma_{g''} \phi_{g''}(\mathbf{x}), \quad (11)$$

where  $\phi_{g''}(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{g''}, \mathbf{P}_{g''})$  is the  $g''$ -th component and  $\sum_{g''=1}^{G''} \gamma_{g''} = 1$ . For notational convenience, the subscript  $k$  and the measurement matrix  $\mathbf{Z}$  in (10) are dropped. Our goal is to approximate  $p(\mathbf{x})$  as a simplified mixture with fewer components in the form of  $\tilde{p}(\mathbf{x}) = \sum_{g=1}^G \gamma_g \phi_g(\mathbf{x})$ , where  $\phi_g(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_g, \mathbf{P}_g)$  with  $\sum_{g=1}^G \gamma_g = 1$  and  $G < G''$ . Given a distance measurement  $D(\phi_{g''}(\mathbf{x}), \phi_g(\mathbf{x}))$  between the functions  $\phi_{g''}(\mathbf{x})$  and  $\phi_g(\mathbf{x})$ , the error of approximating  $\phi_{g''}(\mathbf{x})$  with  $\phi_g(\mathbf{x})$  is given by  $D(\phi_{g''}(\mathbf{x}), \phi_g(\mathbf{x})) = (\int (\phi_{g''}(\mathbf{x}) - \phi_g(\mathbf{x}))^2 d\mathbf{x})^{1/2}$ .

Traditionally, the simplification is done by resampling followed by clustering using the K-means or EM algorithm. However, the complexity depends exponentially on the dimension of the state and hence it will soon become infeasible. As an alternative, simpler algorithms such as pruning and merging can be used for mixture reduction. In the pruning method, GM components with relatively lower weights are discarded directly and then the remaining components are renormalized for the estimation. Rather than simple pruning, the merging method performs mixture reduction by calculating the expected parameters for each merged component across a set of components to be fused. It has been proved that pruning is inferior to the merging method [39]. Moreover, for better determination of the components to be merged, a Kullback-Leibler-based approach is proposed in [40] where mixture reduction is achieved by minimizing an upper bound on the increase of the Kullback-Leibler divergence between the original and reduced mixtures. In this paper, instead of using the pruning and merging methods, we follow our preliminary work in [28] to adopt a two-step algorithm in [41] for model order reduction, which offers more accurate estimates than the Kullback-Leibler approach [40] at the cost of a small increase of computational complexity for mixture reduction.

Due to space limitations, interested readers are referred to Part 1 in the supplementary material for a detailed comparison of the two-step algorithm against the other approaches. The two-step algorithm goes as follows:

At the  $t$ -th iteration, the component mixture is partitioned into  $G$  groups  $\{S_1^t, S_2^t, \dots, S_G^t\}$ .

- Mean update: The representative component  $C_g^{(t)}(\mathbf{x})$  for  $S_g^t$  that minimizes the local quantization error of the group is solved using coordinate descent,  $C_g^{(t)}(\mathbf{x}) = \arg \min_{C(\mathbf{x})} \sum_{g'' \in S_g^t} \gamma_{g''} \int (C(\mathbf{x}) - \phi_{g''}(\mathbf{x}))^2 d\mathbf{x}$ .
- Clustering: Given  $C^{(t)} = \{C_g^{(t)}(\mathbf{x})\}_{g=1}^G$ , one re-assigns  $\phi_{g''}(\mathbf{x})$  to the nearest  $C_g^{(t)}(\mathbf{x})$  based on the distortion measure  $D_{g,g''} = D(C_g^{(t)}(\mathbf{x}), \phi_{g''}(\mathbf{x}))$ , and then updates  $S_g^t$ .

The above process is repeated until either the change in total distortion or  $C^{(t)} = \{C_g^{(t)}(\mathbf{x})\}_{g=1}^G$  is less than a certain threshold, or a maximum number of iterations is reached. The main steps of the GM simplification algorithm are summarized in Algorithm 2, where  $k(r)$  and  $k'(r)$  are Gaussian kernels,  $r_{gg''}^{p,q} = (\boldsymbol{\mu}_g^p - \boldsymbol{\mu}_{g''}^q)^T (\mathbf{P}_g^q - \mathbf{P}_{g''}^q)^{-1} (\boldsymbol{\mu}_g^p - \boldsymbol{\mu}_{g''}^q)$ ,  $c_K^{-1}$  is a normalized constant,  $\varepsilon_1$  and  $\varepsilon_2$  are stopping thresholds for the update of  $\boldsymbol{\mu}_g$  and  $\mathbf{P}_g$ , respectively.

It is seen that by using this GM simplification algorithm, the mixture model in (10) can be approximated by the mixture model in (7) with fewer components, so that the number of components after each iteration can be maintained at a constant level. If the dimension of the state is  $d$ , the complexity of the above-mentioned two-step algorithm is  $\mathcal{O}(T(L+G)G^2d^3)$ , where  $T$  and  $L$  are respectively the numbers of iterations for the outermost loop and the loop from line 8 to 21 in Algorithm 2. Their values are typically small<sup>4</sup> [41].

### C. ADAPTIVE FADING BUKF (AF-BUKF)

As mentioned earlier, the BUKF-SGM algorithm works well when the system dynamics and measurement relationships are known *a priori*. However, its performance may be degraded when the parameters of the state and measurement noises are not known exactly or time-varying. In order to achieve better state estimation, we propose in this subsection an adaptive fading BUKF algorithm to adapt the noise parameters. We first introduce a novel adaptive fading UKF (AF-UKF) that can adaptively adjust the state and measurement noise covariance in the conventional UKF to better explain the observations. Then, it is extended to the BUKF framework.

Given the residual in the measurement update of UKF (line 29 of Algorithm 1), the measurement residual covariance can be calculated as

$$\mathbf{P}_k^e = \mathbb{E} \left[ \mathbf{e}_k \mathbf{e}_k^T \right] = \tilde{\mathbf{P}}_{z,k|k-1} + \tilde{\mathbf{R}}_k, \quad (12)$$

<sup>4</sup>Preliminary results in [33] has demonstrated the potential of the proposed BUKF approach over other methods in a simulated example.  $T$  is around 15,  $L$  is around 6 as observed in our experiments.

### Algorithm 2 The Two-Step Order Reduction Algorithm

```

1: Input:  $\tilde{p}(\mathbf{x}) = \sum_{g''=1}^{G''} \gamma_{g''} N(\mathbf{x}, \boldsymbol{\mu}_{g''}, \mathbf{P}_{g''})$ , partitioning
 $\phi_{g''}(\mathbf{x})$  into  $\{S_1, S_2, \dots, S_G\}$ ;  $t = 0$ ,  $e^{(t)} = 10^{10}$ ,
 $\varepsilon_1 = 0.01$ ,  $\varepsilon_2 = 0.01$ ;
2: repeat
3:    $t = t + 1$ ;
4:   Mean update
5:   for  $g = 1, 2, \dots, G$  do
6:      $\boldsymbol{\mu}_g^0 = \frac{\sum_{g'' \in S_g} \gamma_{g''} \boldsymbol{\mu}_{g''}}{\sum_{g'' \in S_g} \gamma_{g''}}$ ;
7:      $\mathbf{P}_g^0 = \frac{\sum_{g'' \in S_g} \gamma_{g''} \left( \mathbf{P}_{g''}^0 + (\boldsymbol{\mu}_g^0 - \boldsymbol{\mu}_{g''}^0)(\boldsymbol{\mu}_g^0 - \boldsymbol{\mu}_{g''}^0)^T \right)}{\sum_{g'' \in S_g} \gamma_{g''}}$ ;
8:     repeat
9:        $p = 0, q = 0$ ;
10:      repeat
11:         $p = p + 1$ ;
12:         $\boldsymbol{\mu}_g^{p+1} = \left( \frac{\sum_{g' \in S_g} \gamma_{g'} k'(r_{gg''}^{p,0}) (\mathbf{P}_{g'}^p - \mathbf{P}_g^0)^{-1}}{|\mathbf{P}_{g'}^p - \mathbf{P}_g^0|^{1/2}} \right)^{-1}$ 
 $\times \sum_{g'' \in S_g} \frac{\gamma_{g''} k'(r_{gg''}^{p,0}) (\mathbf{P}_{g''}^p - \mathbf{P}_g^0)^{-1} \boldsymbol{\mu}_{g''}^p}{|\mathbf{P}_{g''}^p - \mathbf{P}_g^0|^{1/2}}$ ;
13:      until  $\left\| \boldsymbol{\mu}_g^p - \boldsymbol{\mu}_g^{p-1} \right\|_F \leq \varepsilon_1$ 
14:      repeat
15:         $\mathbf{U} = (\boldsymbol{\mu}_{g''}^p - \boldsymbol{\mu}_g^p) (\boldsymbol{\mu}_{g''}^p - \boldsymbol{\mu}_g^p)^T$ ;
16:         $\mathbf{M} = k(r_{gg''}^{p,q}) \mathbf{P}_{g''}^q - 4k'(r_{gg''}^{p,q}) \mathbf{U}$ 
 $- (\mathbf{P}_g^q + \mathbf{P}_{g''}^q)^{-1} \mathbf{P}_{g''}^q$ ;
17:         $\mathbf{P}_g^{q+1} = \left( \sum_{g'' \in S_g} \frac{(\mathbf{P}_g^q + \mathbf{P}_{g''}^q)^{-1} \gamma_{g''} k(r_{gg''}^{p,p})}{|\mathbf{P}_g^q + \mathbf{P}_{g''}^q|^{1/2}} \right)^{-1}$ 
 $\times \sum_{g'' \in S_g} \frac{(\mathbf{P}_g^q + \mathbf{P}_{g''}^q)^{-1} \gamma_{g''}}{|\mathbf{P}_g^q + \mathbf{P}_{g''}^q|^{1/2}} \mathbf{M}$ ;
18:         $q = q + 1$ ;
19:      until  $\left\| \mathbf{P}_g^q - \mathbf{P}_g^{(q-1)} \right\|_F \leq \varepsilon_2$ 
20:       $\gamma_g = |2\mathbf{P}_g^q|^{1/2} \sum_{g'' \in S_g} \frac{\gamma_{g''} k'(r_{gg''}^{p,p})}{|\mathbf{P}_g^{q+1} + \mathbf{P}_{g''}^q|^{1/2}}$ ;
21:      until  $p \leq 1$  and  $q \leq 1$ 
22:    end for  $g$ 
23:    Clustering
24:     $e^{(t)} = 0$ 
25:    for  $g'' = 1, 2, \dots, G''$  do
26:      for  $g = 1, 2, \dots, G$  do
27:         $\mathbf{Z} = \sum_{g'' \in S_g} \gamma_{g''}$ ;
28:         $D(\phi_{g''}, C_g) = c_K^{-1} \left( |2\mathbf{P}_{g''}^q|^{1/2} + \frac{\gamma_g^2}{Z_i^2} |2\mathbf{P}_g^q|^{1/2} \right.$ 
 $\left. - 2 \frac{\gamma_g}{Z_i} k(r_{gg''}^{p,q}) |\mathbf{P}_g^q + \mathbf{P}_{g''}^q|^{1/2} \right)$ ;
29:      end for
30:      Assign  $\phi_{g''}$  to  $S_g$ :  $g = \arg \min_k D(\phi_{g''}, C_k)$ ;
31:       $e^{(t)} = e^{(t)} + D(\phi_{g''}, C_k)$ ;
32:    end for
33:    until  $|e^{(t)} - e^{(t-1)}| \leq 0.001 |e^{(t)}|$  or  $t == t_{max}$ 
34:  Output:  $p(\mathbf{x}) = \sum_{g=1}^G \gamma_g \phi_g(\mathbf{x}) = \sum_{g=1}^G \gamma_g N(\mathbf{x}; \boldsymbol{\mu}_g, \mathbf{P}_g)$ .

```

where  $\tilde{\mathbf{R}}_k$  is the measurement error covariance resulted from the true state error covariance  $\tilde{\mathbf{P}}_{x,k|k-1}$ , and  $\tilde{\mathbf{R}}_k$  is the true measurement noise covariance.  $\tilde{\mathbf{P}}_{zz,k|k-1}$ ,  $\tilde{\mathbf{P}}_{x,k|k-1}$  and  $\tilde{\mathbf{R}}_k$  are different from their nominal values denoted by  $\mathbf{P}_{zz,k|k-1}$ ,  $\mathbf{P}_{x,k|k-1}$  and  $\mathbf{R}_k$ , respectively.

We aim to find positive scalar correction or fading factors  $\rho_{s,k}$  and  $\rho_{m,k}$  for  $\mathbf{P}_{x,k|k-1}$  and  $\mathbf{R}_k$ , respectively, such that

$$\tilde{\mathbf{P}}_{x,k|k-1} = \rho_{s,k} \mathbf{P}_{x,k|k-1}, \quad (13a)$$

$$\tilde{\mathbf{R}}_k = \rho_{m,k} \mathbf{R}_k. \quad (13b)$$

In other words, we hope to compensate for the mis-specified covariance by adaptively scaling the covariance matrices through the fading factors, while the nominal shape of the covariance is still fixed. Since  $\tilde{\mathbf{P}}_{zz,k|k-1}$  and  $\mathbf{P}_{zz,k|k-1}$  can be approximated by the unscented transformation of  $\tilde{\mathbf{P}}_{x,k|k-1}$  and  $\mathbf{P}_{x,k|k-1}$ , the ratio of the latter quantities, i.e.,  $\rho_{s,k}$  in (13a), is approximately equal to that of the former. Therefore, from (13a), we have

$$\tilde{\mathbf{P}}_{zz,k|k-1} \approx \rho_{s,k} \mathbf{P}_{zz,k|k-1}. \quad (14)$$

Substituting (13b) and (14) into (12) gives

$$\mathbf{P}_k^e = \rho_{s,k} \mathbf{P}_{zz,k|k-1} + \rho_{m,k} \mathbf{R}_k. \quad (15)$$

Instead of matching the whole matrix, we take the trace of the left and right hand sides of (15), yielding the following equation at each time instant  $k$ :

$$\text{tr}(\mathbf{P}_k^e) = \rho_{s,k} \text{tr}(\mathbf{P}_{zz,k|k-1}) + \rho_{m,k} \text{tr}(\mathbf{R}_k). \quad (16)$$

Note that the measurement residual covariance  $\mathbf{P}_k^e$  can be estimated empirically as:

$$\hat{\mathbf{P}}_k^e = \frac{1}{M} \sum_{t=k-M+1}^k \mathbf{e}_t \mathbf{e}_t^T, \quad (17)$$

where  $M$  specifies the window size for estimating the empirical covariance. Thus, the estimation of the fading factors reduces to a recursive linear regression problem over time as

$$\text{tr}(\hat{\mathbf{P}}_k^e) \approx \rho_{s,k} \text{tr}(\mathbf{P}_{zz,k|k-1}) + \rho_{m,k} \text{tr}(\mathbf{R}_k), \quad (18)$$

where  $k = 1, 2, \dots$ ,  $\rho_{s,U} \geq \rho_{s,k} \geq \rho_{s,L} \geq 0$ ,  $\rho_{m,U} \geq \rho_{m,k} \geq \rho_{m,L} \geq 0$ , and  $\rho_{s,U}$  and  $\rho_{s,L}$  are respectively the upper and lower bounds of  $\rho_{s,k}$  (likewise for  $\rho_{m,k}$ ).

To solve the above linear regression problem recursively, one can use KF with the following state-space model:

$$\mathbf{x}_k^c = \mathbf{F}_k^c \mathbf{x}_{k-1}^c + \boldsymbol{\xi}_k, \quad (19a)$$

$$\mathbf{y}_k^c = \mathbf{H}_k^c \mathbf{x}_k^c + \boldsymbol{\eta}_k, \quad (19b)$$

where  $\mathbf{x}_k^c = [\rho_{s,k}, \rho_{m,k}]^T$  stands for the state vector containing the fading factors,  $\mathbf{y}_k^c = \text{tr}(\hat{\mathbf{P}}_k^e)$ ,  $\mathbf{F}_k^c = \mathbf{I}$ ,  $\mathbf{H}_k^c = [\text{tr}(\mathbf{P}_{zz,k|k-1}), \text{tr}(\mathbf{R}_k)]$ , and  $\boldsymbol{\xi}_k$  and  $\boldsymbol{\eta}_k$  are the corresponding state and measurement Gaussian distributed noises, respectively. This allows the fading factors to be estimated recursively using Kalman filtering approaches. To be more specific, if  $\boldsymbol{\xi}_k$  and  $\boldsymbol{\eta}_k$  are modeled by single Gaussian components as  $\boldsymbol{\xi}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k^c)$  and  $\boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k^c)$ , the estimation of fading factors can be achieved by using the conventional

### Algorithm 3 The Proposed AF-BUKF/S Algorithm

---

```

1: AF-BUKF Tracker
2: Initialization:  $p(\mathbf{x}_0) = \sum_{g=1}^G \gamma_{g,0} \mathcal{N}(\mathbf{x}_0; \boldsymbol{\mu}_{g,0}, \mathbf{P}_{g,0})$ ;
3: for  $k = 1, 2, \dots$  do
4:   Time Update
5:   for  $i = 1 : I$  do
6:     for  $g = 1 : G$  do
7:        $g' = g + (i - 1)G$ ;  $\gamma_{g',k|k-1} = \gamma_{g,k-1} \alpha_{i,k}$ ;
8:       Estimate  $\boldsymbol{\mu}_{g',k|k-1}$  and  $\mathbf{P}_{g',k|k-1}$  from  $\boldsymbol{\mu}_{g,k-1}$ ,
           $\mathbf{P}_{g,k-1}$  and  $\mathbf{Q}_{i,k}$  (UKF Time Update);
9:     end for  $g$ 
10:   end for  $i$ 
11:   Covariance Compensation
12:    $g'_f = \max_{g'} \{\gamma_{g',k|k-1}\}_{g'=1:GI}$ ;
13:   Estimate  $\boldsymbol{\mu}_{z,k|k-1}$  and  $\mathbf{P}_{zz,k|k-1}$  with  $\boldsymbol{\mu}_{g'_f,k|k-1}$  and
           $\mathbf{P}_{g'_f,k|k-1}$  using Unscented Transform;
14:    $\mathbf{e}_k = \mathbf{z}_k - \boldsymbol{\mu}_{z,k|k-1}$ ,  $\hat{\mathbf{P}}_k^e = \frac{1}{M} \sum_{t=k-M+1}^k \mathbf{e}_t \mathbf{e}_t^T$ ;
15:   Compute  $\rho_{s,k}$  and  $\rho_{m,k}$  by solving
           $\text{tr}(\hat{\mathbf{P}}_k^e) \approx \rho_{s,k} \text{tr}(\mathbf{P}_{zz,k|k-1}) + \rho_{m,k} \text{tr}(\mathbf{R}_{1,k})$ ;
16:   Set  $g'_f - th$  state error covariance to  $\rho_{s,k} \mathbf{P}_{g'_f,k|k-1}$ ;
17:   Measurement Update
18:   for  $j = 1 : J$  do
19:     for  $g' = 1 : GI$  do
20:        $g'' = g' + (j - 1)GI$ ;  $\gamma_{g'',k|k-1} = \gamma_{g,k-1} \alpha_{i,k}$ ;
21:       if  $g'' == g'_f$  then
22:         Estimate  $\boldsymbol{\mu}_{g'',k|k-1}$  and  $\mathbf{P}_{g'',k|k-1}$  with
           $\boldsymbol{\mu}_{g',k-1}$ ,  $\mathbf{P}_{g',k-1}$  and  $\rho_{m,k} \mathbf{R}_{j,k}$ 
          (UKF Measurement Update);
23:       else
24:         Estimate  $\boldsymbol{\mu}_{g'',k|k-1}$  and  $\mathbf{P}_{g'',k|k-1}$  with
           $\boldsymbol{\mu}_{g',k-1}$ ,  $\mathbf{P}_{g',k-1}$  and  $\mathbf{R}_{j,k}$ 
          (UKF Measurement Update);
25:       end if
26:        $\mathbf{e}_k = \mathbf{z}_k - h_k(\boldsymbol{\mu}_{g'',k|k-1}) - \bar{v}$ ;
27:        $\boldsymbol{\mu}_{g'',k} = \boldsymbol{\mu}_{g'',k|k-1} + \mathbf{K}_{g'',k} \mathbf{e}_k$ ;
28:        $\mathbf{P}_{g'',k} = (\mathbf{I} - \mathbf{K}_{g'',k} \mathbf{H}_{g'',k}) \mathbf{P}_{g'',k|k-1}$ ;
29:        $\gamma_{g'',k}^* = \gamma_{g'',k|k-1} \beta_{j,k} p_j(\mathbf{z}_k | \boldsymbol{\mu}_{g'',k})$ ;
30:     end for  $g'$ 
31:   end for  $j$ 
32:    $\gamma_{g'',k} = \frac{\gamma_{g'',k}^*}{\sum_{g''=1}^{G'} \gamma_{g'',k}^*}$ ,  $\hat{\mathbf{x}}_{k|k} = \sum_{g''=1}^{G'} \gamma_{g'',k} \boldsymbol{\mu}_{g'',k}$ ;
33:   Order Reduction
34:   Compute  $\{\gamma_{g,k}\}_{g=1}^G$ ,  $\{\boldsymbol{\mu}_{g,k}\}_{g=1}^G$  and  $\{\mathbf{P}_{g,k}\}_{g=1}^G$ .
35: end for  $k$ 
36: Backward Smoother
37: for  $k = N - 1, \dots, 1$  do
38:   for  $g = 1 : G$  do
39:     Get component indices  $g^*$  and  $i^*$ ;
40:     Compute  $\boldsymbol{\mu}_{g^*,k+1|k}$ ,  $\mathbf{P}_{g^*,k+1|k}$  and  $\mathbf{C}_{g^*,k+1}$ ;
41:      $\mathbf{K}_{g,(k|N)} = \mathbf{C}_{g^*,k+1} \mathbf{P}_{g^*,(k+1|k)}^{-1}$ ;
42:      $\boldsymbol{\mu}_{g,(k|N)} = \boldsymbol{\mu}_{g^*,k} + \mathbf{K}_{g,(k|N)} (\boldsymbol{\mu}_{g^*,(k+1|N)} - \boldsymbol{\mu}_{g^*,(k+1|k)})$ ;
43:      $\mathbf{P}_{g,(k|N)} = \mathbf{P}_{g^*,k} + \mathbf{K}_{g,(k|N)} (\mathbf{P}_{g^*,(k+1|N)} - \mathbf{P}_{g^*,(k+1|k)}) \mathbf{K}_{g,(k|N)}^T$ ;
44:   end for  $g$ 
45:    $\hat{\mathbf{x}}_{(k|N)} = \sum_{g=1}^G \gamma_{g,(k|N)} \boldsymbol{\mu}_{g,(k|N)}$ ;
46: end for  $k$ 

```

---

KF. Particularly, the parameters for  $\mathbf{Q}_k^c$  and  $\mathbf{R}_k^c$  should be set according to the characteristics of systems to be estimated. More precisely, let  $\mathbf{Q}_k^c = \text{diag}(a, b)$  where  $\text{diag}(a, b)$  represents the diagonal matrix with diagonal elements  $a$  and  $b$ . The setting of values for  $a$  and  $b$  should be as follows: for systems with inaccurate state equations, a relatively larger value shall be assigned to  $a$  so that more compensation for the state can be achieved than that for measurement, while  $b$  shall be larger than  $a$  for systems with inaccurate measurement equations. On the other hand, we also examine the case where  $\xi_k$  is modeled as GMs as in (6a). This allows faster tracking of the fading factors and the BKF algorithm [28]–[30] we developed previously can be utilized. The advantages of GM modeling for fading factors over the single Gaussian representation are demonstrated with simulated data and UASs examples in Part 2 of the supplementary material. Moreover, in each recursion, the estimated parameters will be projected to the bound constraints above. This is in contrast to other adaptive KFs [34]–[37], [42]–[45]. Due to page limitations, a comparison of the proposed AF-UKF and the state-of-the-art approaches in [34]–[37] is given in the supplementary materials. It is shown that the proposed AF-UKF offers a better performance than the conventional methods. It should be noted that the measurement covariance matrix can also be estimated using variational Bayesian approach in [38]. However, its extension to uncertainty in the state noise is not yet available. On the other hand, the proposed AF approach has fewer parameters to be estimated and is applicable to both state and measurement noises. Therefore, we shall focus on the AF approach in this paper.

To incorporate the proposed AF-UKF into the BUKF framework, one can notice that the basic UKF is used to compute each of the  $G''$  components in (10) before they are simplified to a GM with  $G$  components in (7) using Algorithm 1. Ideally, one can simply replace the UKF in the AF-BUKF with the AF-UKF algorithm. However, such adaptation may result in all mixture components of the state and measurement noise converging to a single solution. In case of sudden changes of state or measurement covariance, these changes can no longer be modeled as a GM using the prior distribution. Therefore, we propose to select the most likely component from the GMs and utilize the AF-UKF instead of the conventional UKF to update its distribution. The AF component together with other components will then be approximated as a GM with  $G$  components. This yields the AF-BUKF tracker in Algorithm 3. Finally, it is worth noting that our AF and SGM approaches are also applicable to other methods for approximating the integrals in (3) and (4) such as cubature KF, EKF and Gauss-Hermite KF. Due to page limitation, the details are omitted.

#### D. ADAPTIVE FADING BUK SMOOTHER (AF-BUKS)

In filtering, the objective is to recursively estimate the current state given the observation up to the current time  $k$ . Smoothing can yield significantly better estimates than filtering if the estimation of the current state can be delayed so that more

data can be utilized in the estimation process [46]. Therefore, it aims to improve the estimation accuracy by utilizing a certain number of future samples in exchange for a delay in obtaining the required smoothed estimates. This is desirable in offline applications or applications which certain delay can be tolerated. Here, we extend the AF-BUKF algorithm by means of the backward smoother, i.e. Rauch-Tung-Striebel Smoother [47] for further improving the estimation accuracy. More specifically, it can be shown in [48] that the desired pdf at the current time  $k$ ,  $p(\mathbf{x}_k|\mathbf{Z}_N)$ , is given by the following general Bayesian smoothing formula:

$$p(\mathbf{x}_k|\mathbf{Z}_N) = p(\mathbf{x}_k|\mathbf{Z}_k) \int \frac{p(\mathbf{x}_{k+1}|\mathbf{x}_k)p(\mathbf{x}_{k+1}|\mathbf{Z}_N)}{p(\mathbf{x}_{k+1}|\mathbf{Z}_k)} d\mathbf{x}_{k+1}, \quad (20)$$

where  $p(\mathbf{x}_{k+1}|\mathbf{Z}_N)$  is the posterior pdf obtained from backward smoothing at time  $k+1$ . Since it has used data up to time  $N$ , we have used  $\mathbf{Z}_N$  in the condition of  $p(\mathbf{x}_{k+1}|\mathbf{Z}_N)$ . It can be seen that the pdf obtained in the Bayesian filtering step,  $p(\mathbf{x}_k|\mathbf{Z}_k)$ , is updated through the integral term in the smoothing step. In the Gaussian case, the integral and the entire computation can be further simplified to the RTS smoother. In preparation for the approximated Bayesian smoothing, we save the  $GII$  components before the mixture simplification at time  $k+1$ :  $\tilde{p}(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1}) = \sum_{g''=1}^{G''} \gamma_{g'',k+1} \mathcal{N}(\mathbf{x}_{k+1}; \boldsymbol{\mu}_{g'',k+1}, \mathbf{P}_{g'',k+1})$ . The number of components in  $p(\mathbf{x}_{k+1}|\mathbf{Z}_N)$  is chosen to be  $G$ , which is the same as that of the simplified density  $p(\mathbf{x}_k|\mathbf{Z}_k)$ . Therefore, the backward smoothed pdf of  $p(\mathbf{x}_{k+1}|\mathbf{Z}_N)$  will take the form  $p(\mathbf{x}_{k+1}|\mathbf{Z}_N) = \sum_{g=1}^G \gamma_{g,k+1|N} \mathcal{N}(\mathbf{x}_{k+1}; \boldsymbol{\mu}_{g,k+1|N}, \mathbf{P}_{g,k+1|N})$ . To simplify the evaluation of the integral in (20), we assume that there is a correspondence between the GMs in  $p(\mathbf{x}_{k+1}|\mathbf{Z}_N)$  and  $p(\mathbf{x}_k|\mathbf{Z}_k)$ . Consequently, the update in (20) can be reduced to the update of individual components of the mixture. Moreover, the RTS smoothing can be applied to each of the components to give  $p(\mathbf{x}_k|\mathbf{Z}_N) = \sum_{g=1}^G \gamma_{g,(k|N)} \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{g,k|N}, \mathbf{P}_{g,k|N})$ , where  $\boldsymbol{\mu}_{g,k|N}$  and  $\mathbf{P}_{g,k|N}$  are the mean and covariance of the GMM, respectively. To find the correspondence between the GMM, we propose to find the closest neighbor of  $\mathcal{N}(\mathbf{x}_{k+1}; \boldsymbol{\mu}_{g,(k+1|N)}, \mathbf{P}_{g,k+1|N})$  in  $\tilde{p}(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1})$ , i.e.  $\mathcal{N}(\mathbf{x}_{k+1}; \boldsymbol{\mu}_{g'',k+1}, \mathbf{P}_{g'',k+1})$  for some  $g''$ , since it is possible to trace back the component in  $\tilde{p}(\mathbf{x}_k|\mathbf{Z}_k)$  which leads to the identified component in  $\tilde{p}(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1})$ . Consequently, RTS algorithm can be applied to these components to update the corresponding component in  $p(\mathbf{x}_k|\mathbf{Z}_k)$ . Thus, the GM of the smoothed pdf  $p(\mathbf{x}_k|\mathbf{Z}_N)$  can be obtained from the updated components by appropriate normalization. Particularly, we propose to use the Bhattacharyya distance [49]:

$$\begin{aligned} & D_B(g_{k+1|N}, g''_{k+1}) \\ &= \frac{1}{8} (\boldsymbol{\mu}_{g,k+1|N}, \boldsymbol{\mu}_{g'',k+1})^T \bar{\mathbf{P}} (\boldsymbol{\mu}_{g,k+1|N}, \boldsymbol{\mu}_{g'',k+1}) \\ & \quad + \frac{1}{2} \ln \left( \frac{\det(\bar{\mathbf{P}})}{\sqrt{\det(\mathbf{P}_{g,k+1|N}) \det(\mathbf{P}_{g'',k+1})}} \right), \quad (21) \end{aligned}$$



to measure the distance between the  $g$ -th component  $\mathcal{N}(\mathbf{x}_{k+1}; \boldsymbol{\mu}_{g,k+1|N}, \mathbf{P}_{g,k+1|N})$  and the  $g''$ -th component in  $\{\mathcal{N}(\mathbf{x}_{k+1}; \boldsymbol{\mu}_{g'',k+1}, \mathbf{P}_{g'',k+1})\}_{g''=1,\dots,GIJ}$ , where  $\bar{\mathbf{P}} = \frac{\mathbf{P}_{g,(k+1|N)} + \mathbf{P}_{g'',k+1}}{2}$ , and  $\det(\cdot)$  denotes the operation of matrix determinant. From the identified component in  $\{\mathcal{N}(\mathbf{x}_{k+1}; \boldsymbol{\mu}_{g'',k+1}, \mathbf{P}_{g'',k+1})\}_{g''=1,\dots,GIJ}$ , we can identify the corresponding component  $g^*$  in  $p(\mathbf{x}_k|\mathbf{Z}_k) = \sum_{g=1}^G \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{g,k}, \mathbf{P}_{g,k})$ , and the corresponding indices  $i^*$ , and  $j^*$  which lead to the  $g''$ -th component above. Then, we can update all the components in  $p(\mathbf{x}_k|\mathbf{Z}_k)$  by performing the unscented RTS smoothing steps for each of the components in  $p(\mathbf{x}_k|\mathbf{Z}_N)$  as follows:

**Step 1:** Generate the sigma points  $\{\boldsymbol{\chi}_{l,k-1}\}$ ,  $l = 1, \dots, L$ , with positive weights  $\{w_l\}$  from the Gaussian approximation of  $p(\mathbf{x}_{k-1})$  with mean  $\boldsymbol{\mu}_{g^*,k}$  and covariance  $\mathbf{P}_{g^*,k}$  as follows

$$\boldsymbol{\chi}_{0,k-1} = \boldsymbol{\mu}_{g^*,k}, \quad w_0 = \frac{\kappa}{L + \kappa}, \quad (22a)$$

$$\boldsymbol{\chi}_{l,k-1} = \boldsymbol{\mu}_{g^*,k} + \left( \sqrt{(L+\kappa)\mathbf{P}_{g^*,k}} \right)_l, \quad w_l = \frac{1}{2(L + \kappa)}, \quad (22b)$$

$$\boldsymbol{\chi}_{L+l,k-1} = \boldsymbol{\mu}_{g^*,k} - \left( \sqrt{(L+\kappa)\mathbf{P}_{g^*,k}} \right)_l, \quad w_{L+l} = w_l, \quad (22c)$$

where  $\kappa$  is the scaling parameter, and  $\left( \sqrt{\mathbf{P}_{g^*,k}} \right)_l$  denotes the  $l$ -th column of the matrix square root of  $\mathbf{P}_{g^*,k}$ .

**Step 2:** Propagate each sigma point through the nonlinear system state function to obtain samples of the prediction density  $p(\mathbf{x}_{k+1}|\mathbf{x}_k)$ :

$$\boldsymbol{\chi}_{l,k+1|k} = f_{k+1}(\boldsymbol{\chi}_{l,k}), \quad l = 0, \dots, 2L. \quad (23)$$

**Step 3:** Approximate the predicted mean  $\boldsymbol{\mu}_{g^*,k+1|k}$ , the predicted covariance  $\mathbf{P}_{g^*,k+1|k}$  and the cross-covariance  $\mathbf{C}_{g^*,k+1|k}$  as follows:

$$\boldsymbol{\mu}_{g^*,k+1|k} = \sum_{l=0}^{2L} w_l \boldsymbol{\chi}_{l,k+1|k}, \quad (24a)$$

$$\mathbf{P}_{g^*,k+1|k} = \sum_{l=0}^{2L} w_l (\boldsymbol{\chi}_{l,k+1|k} - \boldsymbol{\mu}_{g^*,k+1|k}) \times (\boldsymbol{\chi}_{l,k+1|k} - \boldsymbol{\mu}_{g^*,k+1|k})^T + \mathbf{Q}_{i^*,k+1}, \quad (24b)$$

$$\mathbf{C}_{g^*,k+1} = \sum_{l=0}^{2L} w_l (\boldsymbol{\chi}_{l,k} - \boldsymbol{\mu}_{g^*,k}) (\boldsymbol{\chi}_{l,k+1|k} - \boldsymbol{\mu}_{g^*,k+1|k})^T, \quad (24c)$$

**Step 4:** Compute the smoother gain  $\mathbf{K}_{g,k|N}$ , the smoothed mean  $\boldsymbol{\mu}_{g,k|N}$  and covariance  $\mathbf{P}_{g,k|N}$ :

$$\mathbf{K}_{g,k|N} = \mathbf{C}_{g^*,k} \mathbf{P}_{g^*,k+1|k}^{-1}, \quad (25a)$$

$$\boldsymbol{\mu}_{g,k|N} = \boldsymbol{\mu}_{g^*,k} + \mathbf{K}_{g,k|N} (\boldsymbol{\mu}_{g,k+1|N} - \boldsymbol{\mu}_{g^*,k+1|k}), \quad (25b)$$

$$\mathbf{P}_{g,k|N} = \mathbf{P}_{g^*,k} + \mathbf{K}_{g,k|N} (\mathbf{P}_{g^*,k+1|N} - \mathbf{P}_{g^*,k+1|k}) \mathbf{K}_{g,k|N}^T. \quad (25c)$$

As such, the smoothed density  $p(\mathbf{x}_k|\mathbf{Z}_N)$  can be obtained by normalizing the weight of the updated components so that they sum to one. The smoothed state (point) estimates is thus the mean of the mixture:

$$\hat{\mathbf{x}}_{k|N} = \sum_{g=1}^G \gamma_{g,k|N} \boldsymbol{\mu}_{g,k|N}. \quad (27)$$

The key ideas and steps of the proposed AF-BUKS algorithm are illustrated by the block diagram in Fig. 1. Finally, the proposed AF-BUKS is summarized in Algorithm 3.

#### IV. APPLICATIONS OF AF-BUKF/S IN UASs AND EXPERIMENTAL RESULTS

In this section, we focus on the application of the proposed AF-BUKF/S algorithms to state estimation in UASs. We shall first describe the state-space model of the UASs used for simulation data generation. Then, the performance of the AF-BUKF/S algorithms are compared with conventional approaches including EKF and UKF on the simulated UAS data. Finally, we present experimental results on real flight data to show the efficiency of the proposed algorithm for state estimation of UASs.

Due to space limitation, comparisons with previous works on adaptive fading are presented in the supplementary materials where we have shown that the proposed AF approach outperforms the state-of-the-arts algorithm in [34]–[37]. We have also compared the proposed algorithm with the STUKF/STCKF algorithms using a maneuvering bearing-only tracking example in [26] and the proposed approach is able to offer better accuracy for state estimation of the moving target. The results are also provided in the supplementary materials.

##### A. STATE-SPACE MODEL OF UASs

The system dynamics of small UASs can be mathematically represented by a continuous-time state-space model as

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) + \boldsymbol{\omega}(t), \quad (28a)$$

$$\mathbf{z}(t) = h(\mathbf{x}(t)) + \mathbf{v}(t), \quad (28b)$$

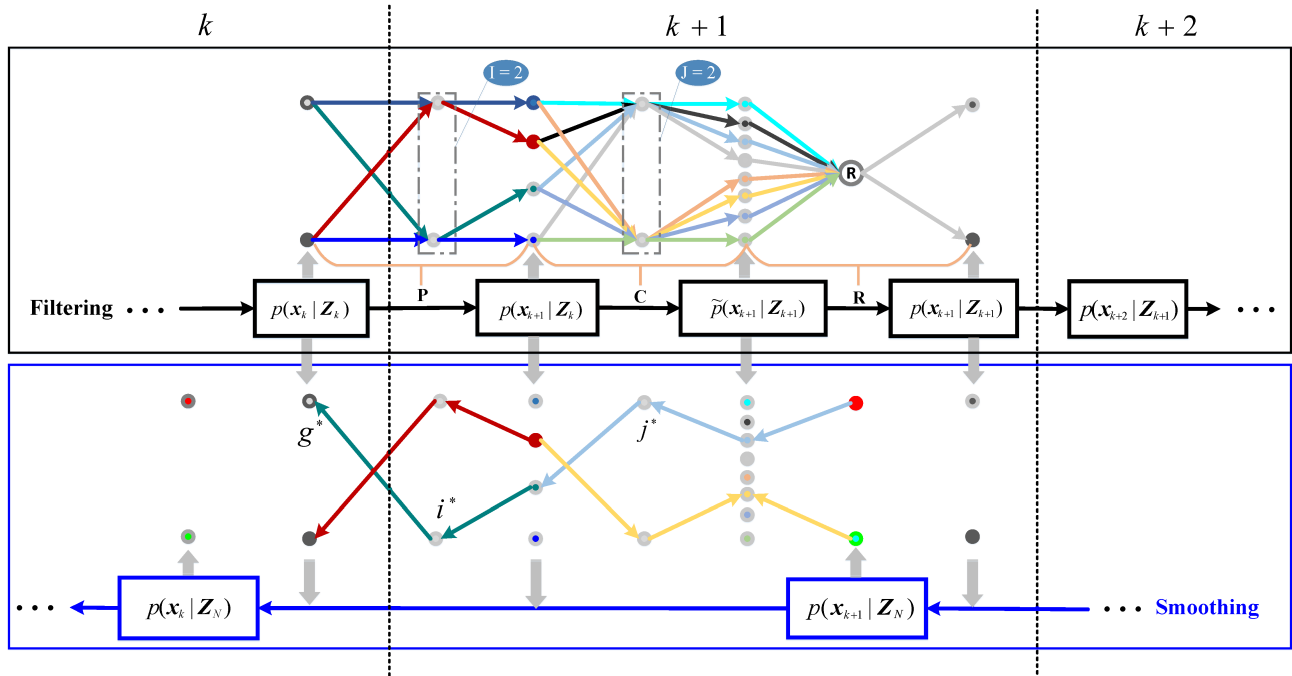
where  $\boldsymbol{\omega}(t)$  and  $\mathbf{v}(t)$  take into account the non-deterministic influences, e.g., state propagation errors and sensor measurement noise that affect the system state dynamics and measurement, respectively. The IMU, GPS and other available measurements can be included into a complete inertial navigation system (INS) for the state estimation. To convert the continuous-time model in (28) to a discrete-time version in (1), the following equation can be used to propagate  $\mathbf{x}_{k-1}$  forward  $\Delta t$  seconds for an estimation of  $\mathbf{x}_k$ :

$$\mathbf{x}_k = \mathbf{x}(k\Delta t) \approx \mathbf{x}_{k-1} + \dot{\mathbf{x}}(t)\Delta t, \quad (29)$$

where  $\Delta t$  is the sampling time.

##### B. SIMULATED DATA CASE

The UASs simulation model designed in [3] is used for performance comparison between the proposed method and the conventional approaches. There are totally 6 simulated UAS datasets in [3] including Letter E (**D1**), Letter E with Gusting (**D2**), Loiter (**D3**), Loiter with Gusting (**D4**), Square Spiral (**D5**) and Square Spiral with Gusting (**D6**). The gusting data (**D2**, **D4** and **D6**) take the practical wind flickering problem into account, as small UAS flying at relatively slow speed may be heavily influenced by the movement of the air



**FIGURE 1.** Illustration for the proposed Bayesian filtering and smoothing. For simplicity, we use  $G = I = J = 2$  for demonstration. It can be seen at time instant  $k + 1$  in the Bayesian filtering, the number of mixture components will grow from 2 to 4 in the prediction (P) step (Time Update) and from 4 to 8 in the subsequent measurement correction (C) step (Measurement Update). Here the two dots in the dotted box represent the  $I = J = 2$  GMs of the state and measurement noises. Using the two-step order reduction (R) algorithm, the component number is reduced to 2. The paths are drawn in different colors for each understanding. For the backward Bayesian smoothing, among components in  $\tilde{p}(x_{k+1} | Z_{k+1})$  (which is saved in the filtering), we first search the closest neighbor for  $\mathcal{N}(x_{k+1}; \mu_{g,(k+1|N)}, P_{g,(k+1|N)})$  in  $p(x_{k+1} | Z_N)$ . Then, by tracing back, say for the component denoted by the red circle, the corresponding component indices  $j^*, i^*$  and  $g^*$  for the component can be identified successively by following the corresponding path in the filtering step. Thus, RTS can be used to estimate the mixture components for  $p(x_k | Z_N)$  in the backward direction.

mass around it. The time duration for the simulation of each data is 400s and the sampling rate is 10Hz.

1) STATE-SPACE MODELS FOR SIMULATION

The nonlinear state and measurement functions are given in (26), as shown at the bottom of the page [3], where the time index is omitted for notation simplicity and the parameters involved are described in Table 1. The continuous time state-space model described above can be linearized and discretized so that an EKF algorithm can be

employed for the state estimation. Due to space limitation, interested readers are referred to [3] for more details of the system model and its EKF implementation. In contrast to the EKF approach, the UKF-based algorithms rely on the unscented transformation and avoid the linearization of the state-space model. Hence, the UKF-based methods can mitigate the error caused by linearization and offer better tracking results as we shall show later in Section IV-B.3. Since the number of states is rather large, the complexity of particle filter-based methods is expected

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \\ \dot{P}_n \\ \dot{P}_e \\ \dot{h}_{MSL} \\ \dot{V}_n \\ \dot{V}_e \\ \dot{V}_d \\ \dot{b}_{\omega_x} \\ \dot{b}_{\omega_y} \\ \dot{b}_{\omega_z} \\ \dot{b}_{ax} \\ \dot{b}_{ay} \\ \dot{b}_{az} \end{bmatrix} = \begin{bmatrix} 1 & \sin \phi \tan \theta & \sin \phi \tan \theta \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi \sec \theta & \cos \phi \sec \theta \\ & & V_n \\ & & V_e \\ & & -V_d \\ & & C_b^{ned} \begin{bmatrix} f_{x,acc} - b_{ax} \\ f_{y,acc} - b_{ay} \\ f_{z,acc} - b_{az} \end{bmatrix} \\ & & 0 \\ & & 0 \\ & & 0 \\ & & 0 \\ & & 0 \\ & & 0 \end{bmatrix} \begin{bmatrix} \omega_{x,gyro} - b_{\omega_x} \\ \omega_{y,gyro} - b_{\omega_y} \\ \omega_{z,gyro} - b_{\omega_z} \end{bmatrix} ; \quad \mathbf{z} = \begin{bmatrix} \psi_{ref} \\ P_{n,GPS} \\ P_{e,GPS} \\ h_{MSL,GPS} \\ V_{n,GPS} \\ V_{e,GPS} \\ V_{d,GPS} \end{bmatrix} = \begin{bmatrix} \psi \\ P_n \\ P_e \\ h_{MSL} \\ V_n \\ V_e \\ V_d \end{bmatrix} \quad (26)$$

TABLE 1. Notations in UASs model.

	Notation	Description	Capture Sensor
$\hat{x}$	$\phi$	Roll attitude	N/A
	$\theta$	Pitch attitude	N/A
	$\psi$	Yaw attitude	N/A
	$\omega$	Rotation rate	Gyroscope
	$b_w$	Gyroscope bias	N/A
	$b_a$	Accelerometer bias	Accelerometer
	$f_{acc}$	Specific force	N/A
	$C_b^{ned}$	Attitude state function	N/A
$z$	$\psi_{ref}$	Reference Yaw	Magnetometer
	$P_{n,GPS}$	North position	GPS
	$P_{e,GPS}$	East position	GPS
	$h_{MSL,GPS}$	Height above mean sea level	GPS
	$V_{n,GPS}$	North velocity	GPS
	$V_{e,GPS}$	East velocity	GPS
	$V_{d,GPS}$	Downward velocity	GPS

Note: N/A denotes 'Not Available'.

to be extremely large. Therefore, they are not included in the comparison.

2) SIMULATION AND PARAMETER SETTINGS

We first describe the settings for simulation data generation. Specifically, the uncertainties of the simulated system were mimicked by the following two scenarios: **C1**) Abrupt change of the measurement noise covariance to illustrate sudden changes in environmental condition<sup>5</sup>: the measurement noise was chosen as  $\bar{R}_k$  in the first 100s and abruptly increased to  $50\bar{R}_k$  in the subsequent time, where  $\bar{R}_k = \text{diag}(0.05^2, 2^2, 2^2, 2^2, 1, 1, 1)$  with  $\text{diag}(a_1, \dots, a_n)$  being a diagonal matrix with diagonal elements  $a_1, \dots, a_n$ ; **C2**) Measurement with impulsive noise components: impulsive noises with occurrence probability of 0.1 were incorporated into the measurement data, and the amplitude is randomly generated with zero mean and variance  $50\bar{R}_k$ .

We now describe the settings for various approaches including EKF, UKF, STUKF, STCKF, BUKG-SGM and AF-BUKF/S. For a fair comparison, the state and measurement covariance matrices of the EKF, UKF, STUKF and STCKF algorithms in our experiments were set identical to those in [3], which are denoted with  $Q_k$  and  $R_k$ , respectively. For BUKF-SGM, the system state noise  $\omega_k$  in (6a) is  $p(\omega_k) = \sum_{i=1}^2 \alpha_{i,k} \mathcal{N}(\omega_k; \bar{\omega}_{i,k}, Q_{i,k})$  with  $\alpha_{1,k} = 0.8$ ,  $\alpha_{2,k} = 0.2$ ,  $\bar{\omega}_{1,k} = \bar{\omega}_{2,k} = 0$ ,  $Q_{1,k} = 10Q_k$  and  $Q_{2,k} = 2Q_k$ . Here, we used a component with relatively larger state covariance to model possible uncertainty stemmed from system state noise, which was manually set by trial and error. The system measurement noise  $v_k$  in (6b) is  $p(v_k) = \sum_{j=1}^2 \beta_{j,k} \mathcal{N}(v_k; \bar{v}_{j,k}, R_{j,k})$  with  $\beta_{1,k} = 0.9$ ,  $\beta_{2,k} = 0.1$ ,  $\bar{v}_{1,k} = \bar{v}_{2,k} = 0$ ,  $R_{1,k} = \bar{R}_k$  and  $R_{2,k} = 10\bar{R}_k$ . The component with larger measurement covariance could be

<sup>5</sup>Since only the measured data and true state are available from [3], only measurement noise mis-matching can be modeled in our simulation.

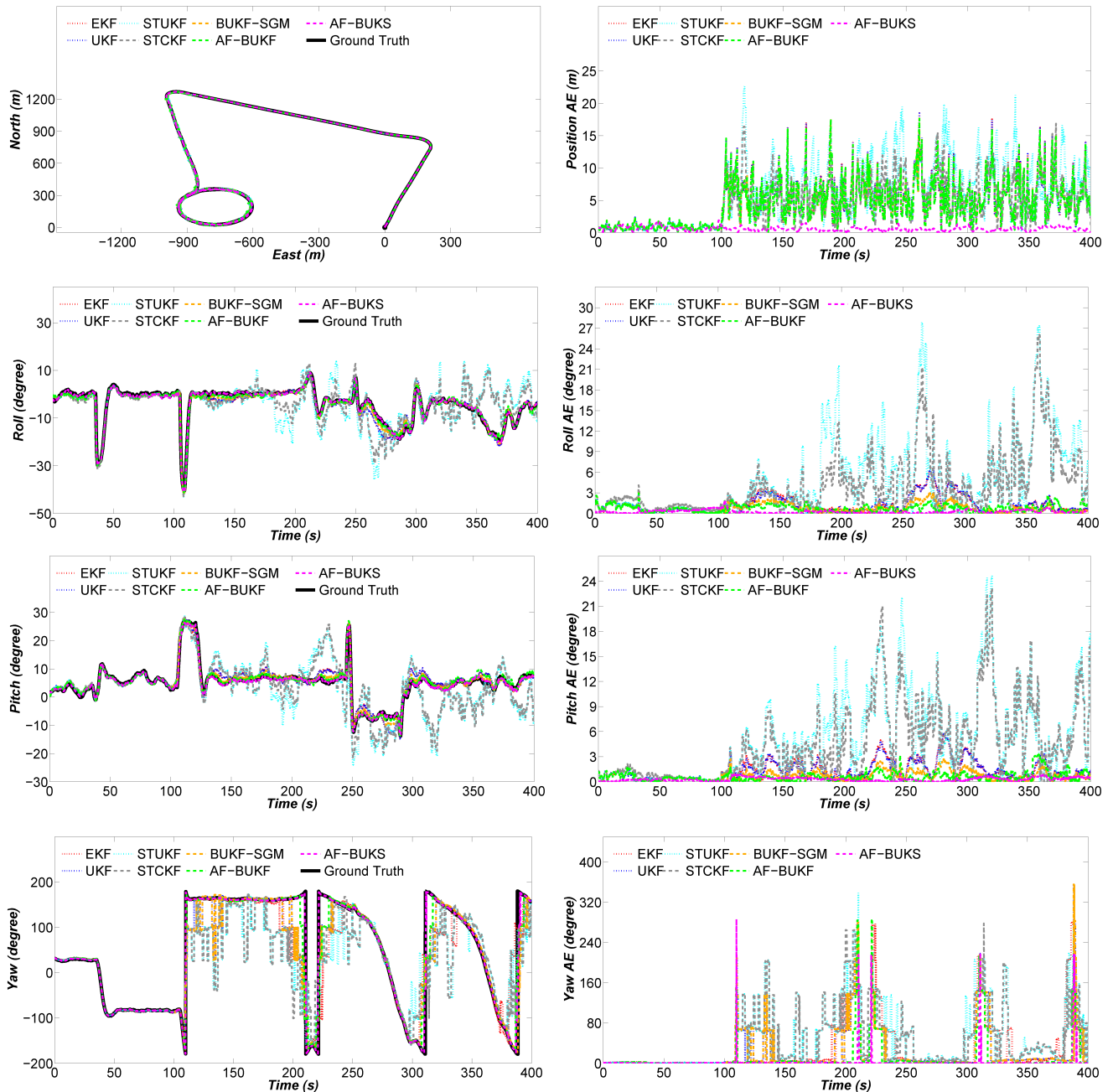
used for impulsive noise suppression. Again the value was manually chosen via trial and error. The initial *a posteriori* pdf of the state was a GM with four components having the same weight.

In order to better show the effectiveness of the covariance compensation that benefits from the incorporated adaptive fading strategy in the proposed AF-BUKF approach, we let one of the GM components, both in state and measurement covariance matrices, be randomly generated with specific variances. Specifically, its settings were the same as those in BUKF-SGM, except that: 1)  $Q_{2,k} = r_s Q_k$ , where the value of  $r_s$  was randomly generated with variance 2; 2)  $R_{2,k} = r_m R_k$ , where the value of  $r_m$  was randomly generated with variance 10; We will see later in Section IV-B3 that the overall performance of the AF-BUKF approach is better than those of EKF and UKF with carefully chosen noises. 3) Additional parameters for the BKF during the estimation of fading factors in the AF-BUKF/S were as follows: the window size was  $M = 20$ ; the initial *a posteriori* pdf of the state vector composed of fading factors was a GM with two components having the same weight; the pdf for the state noise was modeled by a GM with two components, i.e.  $p(\xi_k) = \sum_{i=1}^2 \alpha_{i,k} \mathcal{N}(\xi_k; 0, Q_{i,k}^c)$  with  $\alpha_{1,k} = 0.5$ ,  $\alpha_{2,k} = 0.5$ ,  $Q_{1,k}^c = \text{diag}(10^{-4}, 10^4)$  and  $Q_{2,k}^c = \text{diag}(10^{-2}, 10^2)$ ; the pdf for measurement noise was modeled by single zero mean Gaussian as  $p(\eta_k) = \mathcal{N}(\eta_k; 0, R_k^c)$  with  $R_k^c = 10^{-2}$ ; the upper and lower bounds of the fading factors are set as  $10 \geq f_{s,k} \geq 1$ ,  $100 \geq f_{m,k} \geq 1$ . 4) the delay for smoothing estimation of the AF-BUKS algorithm was set to be 10s (100 samples) which is quite less than the duration for the whole simulation, 400s. This is designed to mimic the case of online tracking when certain delay can be tolerated by the system. Settings for various algorithms were identical in both of the above-mentioned cases.

3) SIMULATION RESULTS

We now compare the performance of the proposed AF-BUKF/S algorithms with the EKF, UKF, STUKF, STCKF and BUKF-SGM approaches in turns through visual and quantitative assessments.

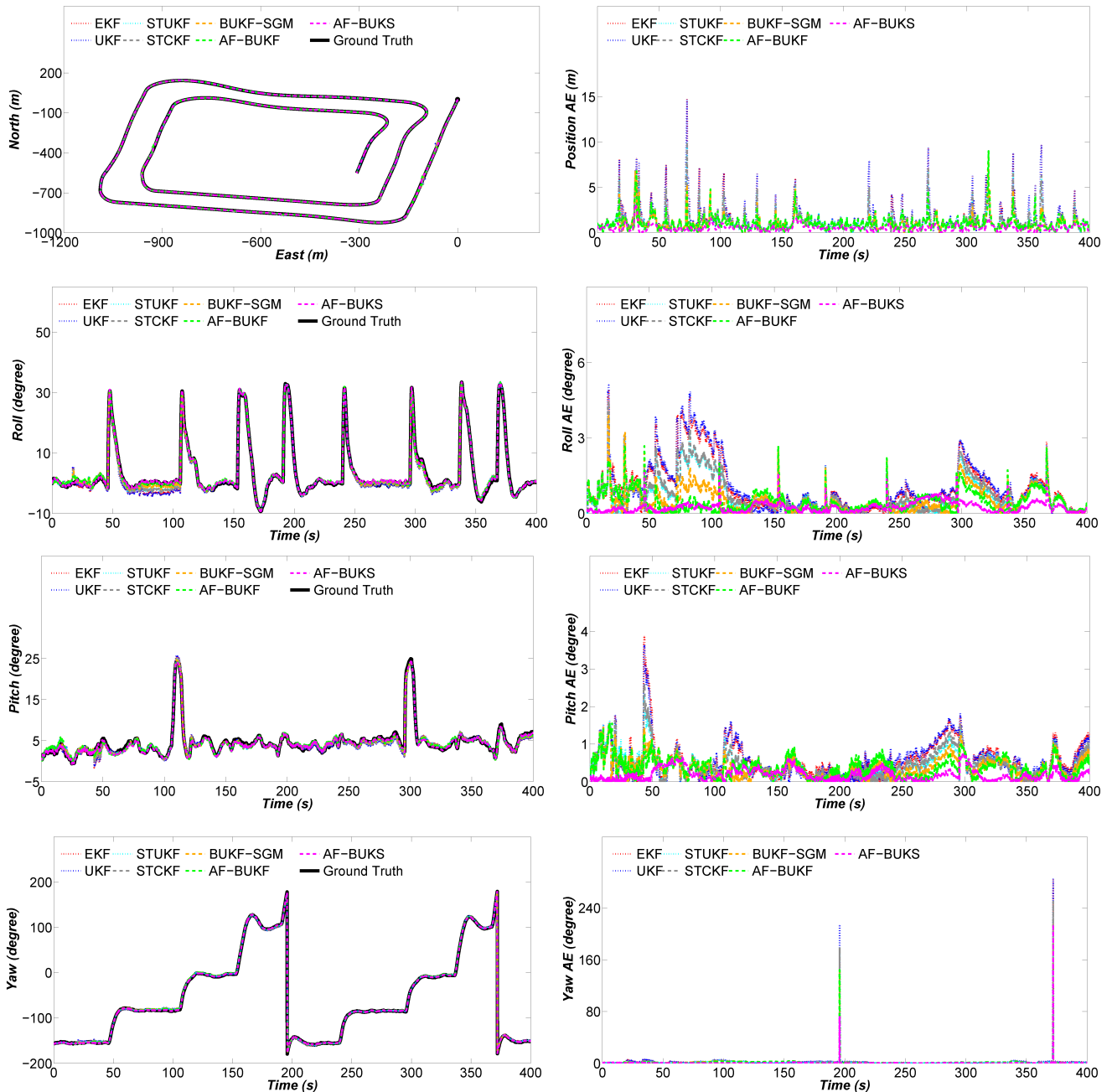
First, the estimated state variables in case **C1** are presented on the left hand side of Fig. 2 for visual assessment. The plotted position (East and North locations) and orientation (Roll, Pitch and Yaw angles) were obtained by averaging the results over 100 Monte Carlo trials of various approaches. Due to page limitations, only the results for **D4** are presented here for visual comparison. We refer interested readers to Part 4 of the supplementary materials for more results. In the first 100s duration for each data, it can be seen that all algorithms can track the state parameters accurately without any obvious visual artifacts, while the AF-BUKS is the best and the rest are with similar performances. However, for the subsequent tracking after the abrupt change in the measurement covariance, the BUKF-based methods perform significantly better than the other approaches, especially for the estimation of roll



**FIGURE 2.** Simulation case C1 for data D4: Left and right columns present the estimated UASs state variables and the corresponding AEs, respectively: the East/North Position, Roll, Pitch and Yaw are shown from the top to the bottom row in order. The position AEs are obtained by averaging the AEs of East and North.

and pitch attitude. The improvement is mainly attributed to their non-Gaussian dynamic state modeling of UASs. This allows them to better track the state variables during rapid system changes by modeling the state and measurement noise as GMs. In contrast, the state or measurement variations in the conventional EKF or UKF are modeled by a single Gaussian component, whose noise covariance may not be exactly known *a priori* especially in time varying and other practical applications. On the other hand, the performances of the STUKF and STCKF algorithms are worse than the

EKF and UKF methods. This suggests that the performance of the Student's *t*-based filters will be degraded substantially when the system noises are not heavy tailed, while the proposed AF-BUKF algorithm does not suffer from such restrictions. Moreover, from the calculated absolute errors (AEs) shown on the right hand side of Fig. 2, it can be seen that the proposed AF-BUKF/S algorithms yield better estimation accuracy than its non-adaptive BUKF-SGM counterpart. This demonstrates the benefits of using the adaptive fading strategy for state estimation with system uncertainties.

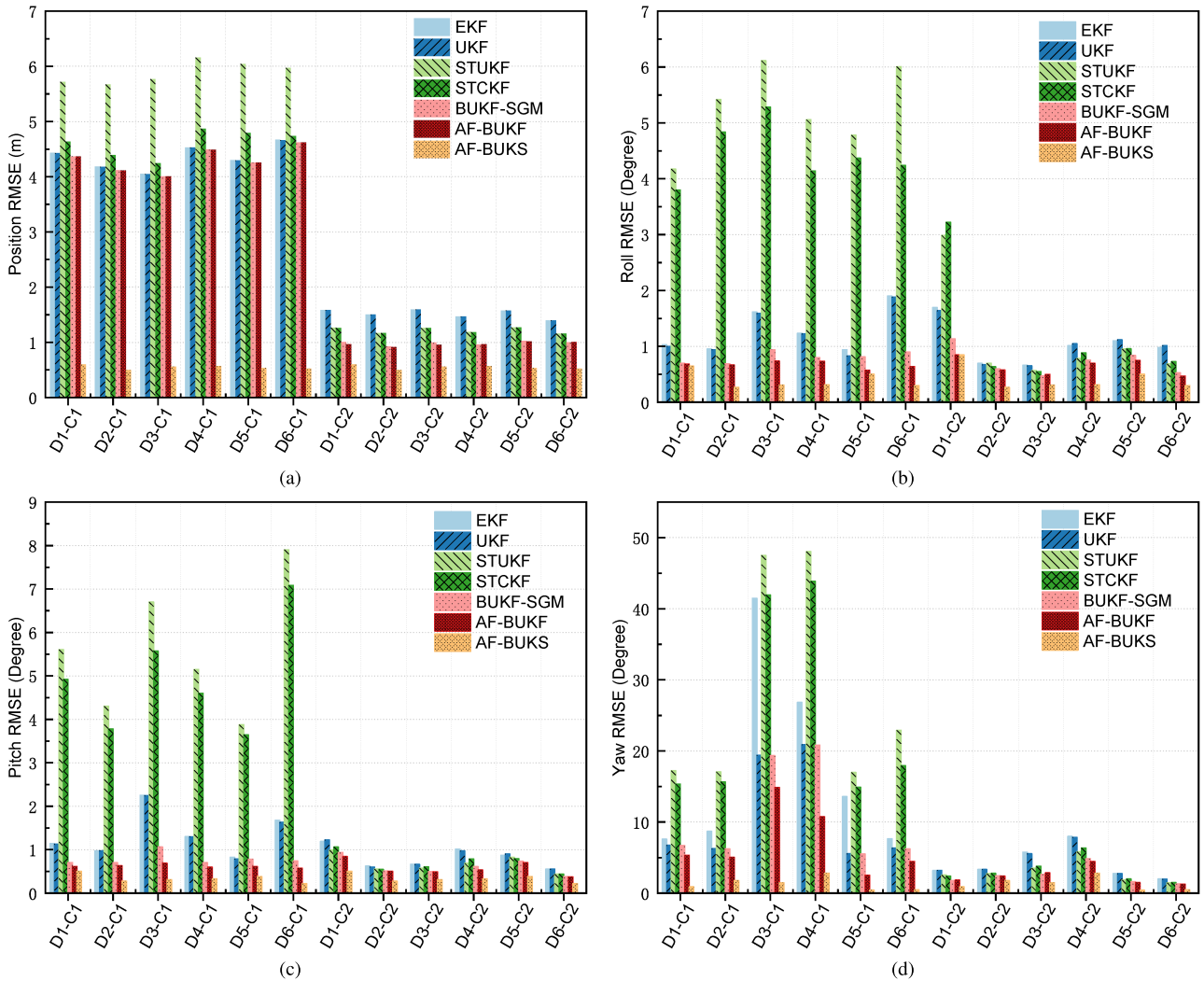


**FIGURE 3.** Simulation case C2 for data D6: Left and right columns present the estimated UASs state variables and the corresponding AEs, respectively: the East/North Position, Roll, Pitch and Yaw are shown from the top to the bottom row in order. The position AEs are obtained by averaging the AEs of East and North.

Furthermore, the AF-BUKS algorithm outperforms the AF-BUKF method on the estimation of various state variables, especially for the position parameter as shown on the top right plot in Fig. 2. We attribute such improvements to the fusion of Bayesian smoothing and adaptive fading strategies. Overall, the proposed AF-BUKF/S methods are particularly useful for UASs state estimation under various uncertainties which may stem from unknown ambient noises such as wind force.

Now, the results of case C2 are demonstrated in Fig. 3. Again, due to space limitations, only the results of D6 are

presented here for illustration and more results can be found in the supplementary materials. To better visualize the different behaviors of various approaches, the locations of impulsive noises are randomly generated and then fixed at some specific time instants and the results with 100 realizations are shown in Fig. 3 for visual assessment. It can be seen that the BUKF-SGM and AF-BUKF/S algorithms are less sensitive to impulsive noises than the other compared approaches. From the AEs presented in the right column in Fig. 3, we can see that the student's  $t$ -based algorithms, STUKF and



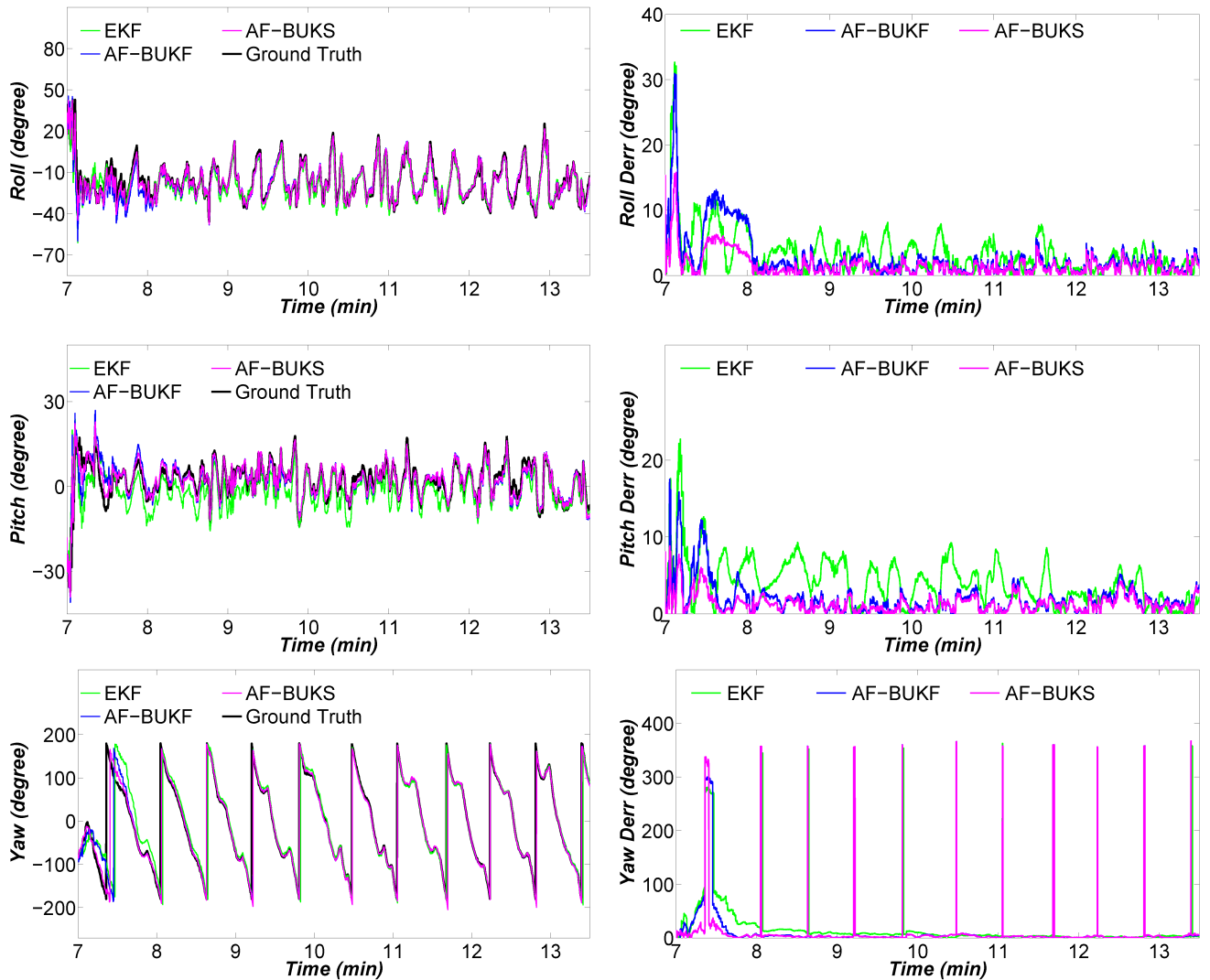
**FIGURE 4.** RMSEs of the UASs state parameters tracked by various algorithms on different datasets (D1-D6) in cases C1 and C2. (a) is for the RMSE of the Position by averaging the RMSEs of East and North. From (b) to (d) are for the RMSE of Roll, Pitch and Yaw, respectively.

STCKF, are slightly better than the EKF/UKF approaches, while the BUKF-based methods can offer significantly better performance than the EKF/UKF algorithms. Moreover, it can also be seen that the overall performance of the AF-BUKF algorithm is significantly better than the BUKF-SGM approach. This is due to the fact that, in our simulation, the amplitude of the impulsive noise was randomly generated with variance  $50\bar{\mathbf{R}}_k$  which is larger than the largest component in the GMs of the measurement noise. In other words, there is a substantial model mismatch in the measurement noise for the BUKF-SGM method. In contrast, the AF-BUKF algorithm is able to automatically compensate for the mismatch through the adaptive fading strategy. Consequently, the adverse effect of the impulsive noise component can be effectively mitigated yielding better performance. Furthermore, the overall performance provided by the AF-BUKS algorithm is better than that of the AF-BUKF. This shows the significance for extending the AF-BUKF to its smoother version via the proposed unscented RTS smoothing concept.

Next, we compare the algorithms quantitatively in terms of the following time-averaged root mean squared error (RMSE), which is defined as

$$RMSE = \frac{1}{K} \sum_{k=1}^K \sqrt{\frac{1}{M} \sum_{m=1}^M (x_k^l(m) - \hat{x}_k^l(m))^2}, \quad (30)$$

where  $x_k^l(m)$  and  $\hat{x}_k^l(m)$  denote respectively the  $l$ -th element of the true and filtered state  $x_k$  and  $\hat{x}$  at the  $m$ -th Monte Carlo run ( $m = 100$  both in cases C1 and C2 for quantitative analysis). It shall be noted that, unlike in the visual analysis, the locations of impulsive noises were randomly generated in the quantitative assessment. The averaged RMSEs for various algorithms on each dataset are compared in Fig. 4. It is seen that the overall performance of the UKF algorithm is slightly better than that of EKF. It seems that the improvement of the unscented transform is somewhat marginal in this case, except for the estimation of the position in case C1 (see left hand sides in Fig. 4(a)). This is mainly due to the mild non-linearity of the simulated UASs. It can also be noticed that



**FIGURE 5.** Experimental results of AF-BUKF/S and EKF algorithms: left column shows the estimated attitude parameters including roll, pitch and yaw; right column presents the corresponding errors. The data length is from 7 to 13.5 minutes, which is the same as in [50].

the BUKF-SGM algorithm yields better performance than the EKF, UKF, STUKF, and STCKF approaches. This further confirms the usefulness of non-Gaussian dynamic state modeling of UASs. Moreover, as seen in Fig. 4, the proposed AF-BUKF/S algorithms outperform the BUKF-SGM and they provide much better performance than the other compared methods. This demonstrates the effectiveness of the proposed AF-BUKF/S algorithms for adaptive estimation of nonlinear states with system uncertainties.

**C. REAL DATA CASE**

To further demonstrate the performance of the proposed AF-BUKF/S algorithms, a benchmark test on real flight data was conducted. We use the real small UAS flight data in [50] to evaluate the performance of the AF-BUKF/S algorithms with comparison against the EKF, UKF, STUKF, STCKF and BUKF-SGM methods. The dataset is obtained from a low-cost IMU/GPS system made by Crossbow. It consists of

GPS data from Micro Nav. and IMU data obtained from gyro and accelerometer readings. The ground truth is obtained by MIDG II, which is an integrated IMU/GPS system equipped on the aircraft that can provide higher quality state variables. The specifications of the MIDG II system are  $0.4^\circ(1\sigma)$  in pitch and roll, and  $2^\circ(1\sigma)$  in heading. The dataset was collected from an unmanned aircraft flying in circular patterns over 13.5 minutes. For detailed description of the nonlinear system models of the flight, we refer interested readers to [50] for more information.

**1) PARAMETER SETTINGS**

For the EKF algorithm, we keep the original settings of the model parameters ( i.e., the initial state and state error covariance, state noise covariance  $Q_k$  and measurement noise covariance  $R_k$ ) in [50] unchanged as they are carefully assigned. Same quantities are used in the UKF, STUKF, STCKF algorithms for fair comparison. For the BUKF-SGM,

**TABLE 2.** MDerr for various algorithms.

Algorithms	Roll (Degree)	Pitch (Degree)	Yaw (Degree)
EKF	3.35	4.16	21.75
UKF	2.98	3.52	18.34
STUKF	3.87	4.88	16.60
STCKF	3.53	4.39	16.17
BUKF-SGM	2.99	2.32	15.26
AF-BUKF	<b>2.56</b>	<b>1.83</b>	<b>12.07</b>
AF-BUKS	<b>2.06</b>	<b>1.54</b>	<b>8.73</b>

the settings are as follows: The system state noise is  $p(\omega_k) = \sum_{i=1}^3 \alpha_{i,k} \mathcal{N}(\omega_k; \bar{\omega}_{i,k}, \mathbf{Q}_{i,k})$  with  $\alpha_{1,k} = 0.5$ ,  $\alpha_{2,k} = 0.25$ ,  $\alpha_{3,k} = 0.25$ ,  $\bar{\omega}_{1,k} = \bar{\omega}_{2,k} = \bar{\omega}_{3,k} = 0$ ,  $\mathbf{Q}_{1,k} = \mathbf{Q}_k$ ,  $\mathbf{Q}_{2,k} = 2\mathbf{Q}_k$  and  $\mathbf{Q}_{3,k} = 5\mathbf{Q}_k$ . The system measurement noise is  $p(\mathbf{v}_k) = \sum_{j=1}^2 \beta_{j,k} \mathcal{N}(\mathbf{v}_k; \bar{\mathbf{v}}_{j,k}, \mathbf{R}_{j,k})$  with  $\beta_{1,k} = 0.5$ ,  $\beta_{2,k} = 0.5$ ,  $\bar{\mathbf{v}}_{1,k} = \bar{\mathbf{v}}_{2,k} = 0$ ,  $\mathbf{R}_{1,k} = \mathbf{R}_k$  and  $\mathbf{R}_{2,k} = 2\mathbf{R}_k$ . The initial *a posteriori* pdf of the state is a GM with four components having the same weight. These parameters were manually selected by trial and error for fair comparison. For the AF-BUKF/S, the settings of the GMs were identical to those in BUKF-SGM. Additional parameters for the BKF during the estimation of fading factors in the AF-BUKF/S are the same as those in Section IV-B.2.

## 2) EXPERIMENTAL RESULTS

We first compare the performance of the proposed AF-BUKF/S algorithms with that of the EKF algorithm in [50]. Specifically, the attitude parameters of roll, pitch and yaw estimated by these two algorithms are presented in the left hand size of Fig. 5. It can be seen that the overall performance of the proposed AF-BUKF algorithm is better than that of the EKF. We attribute this improvement to the non-Gaussian modeling of the state noise of the UASs in the AF-BUKF. This yields more accurate estimation results than the conventional EKF approach with a single component. Next, the error statistics for the AF-BUKF/S and EKF algorithms are plotted in the right hand size of Fig. 5 for comparison. The Euclidean distance between the estimated state  $\hat{\mathbf{x}}_k$  and the ground truth  $\mathbf{x}_k$  is used as the error measure

$$Derr(\hat{\mathbf{x}}_k, \mathbf{x}_k) = \|\hat{\mathbf{x}}_k - \mathbf{x}_k\|_2^2. \quad (31)$$

It can be seen that the AF-BUKS algorithm outperforms the AF-BUKF method, whole performance is much better than the EKF approach. As mentioned earlier, the AF-BUKS algorithm is particularly useful for offline applications or online applications where certain delay can be tolerated. This provides more freedom to users according to the system requirements.

Finally, in Table 2, the proposed AF-BUKF/S algorithms are compared with other approaches based on the mean Euclidean distance criterion, i.e., the errors of (31) averaged over time:

$$MDerr = \frac{1}{N} \sum_{k=1}^N \|\hat{\mathbf{x}}_k - \mathbf{x}_k\|_2^2. \quad (32)$$

The results indicate that the AF-BUKF/S algorithms are able to obtain better estimation accuracy than other approaches.

## V. CONCLUSION

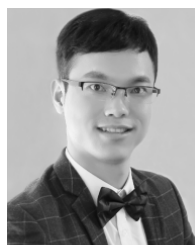
A novel AF-BUKF algorithm has been presented for state estimation of nonlinear and non-Gaussian processes. In contrast to the UKF approach, the AF-BUKF algorithm models the state and noise densities with finite Gaussian mixtures so that the problems caused by system noise uncertainties can be alleviated. Moreover, the AF-BUKF approach is capable of adapting the Gaussian components to further mitigate the adverse effect of inexact knowledge of the state and measurement noise covariance, and thus avoids tedious tuning and offers better adaptation to time-varying environment. In addition, the proposed AF-BUKF algorithm has been further extended to an AF-BUK smoother (AF-BUKS) using the concept of Rauch-Tung-Striebel smoothing for more accurate state estimation. Experimental results on both simulated and real flight data show that the proposed AF-BUKF/S algorithms can better adapt to the uncertainties of system noises and lead to more accurate and reliable estimation of the state of UASs than conventional approaches.

## REFERENCES

- [1] J. Elston, B. Argrow, M. Stachura, D. Weibel, D. Lawrence, and D. Pope, "Overview of small fixed-wing unmanned aircraft for meteorological sampling," *J. Atmos. Ocean. Technol.*, vol. 32, no. 1, pp. 97–115, Jan. 2015.
- [2] T. Gaspar, P. Oliveira, and C. Silvestre, "Model-based filters for 3-D positioning of marine mammals using AHRS- and GPS-equipped UAVs," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 4, pp. 3307–3320, Oct. 2015.
- [3] J. D. Barton, "Fundamentals of small unmanned aircraft flight," *Johns Hopkins APL Tech. Dig.*, vol. 31, no. 2, pp. 132–149, 2012.
- [4] R. W. Beard and T. W. McLain, *Small Unmanned Aircraft: Theory and Practice*. Princeton, NJ, USA: Princeton Univ. Press, 2012.
- [5] S. T. G. Maguire and P. A. Robertson, "UAV attitude estimation using low-frequency radio polarization measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 1, pp. 2–11, Feb. 2017.
- [6] N. Nigam, S. Bieniawski, I. Kroo, and J. Vian, "Control of multiple UAVs for persistent surveillance: Algorithm and flight test results," *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 5, pp. 1236–1251, Sep. 2012.
- [7] J.-S. Lee and K.-H. Yu, "Optimal path planning of solar-powered UAV using gravitational potential energy," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1442–1451, Jun. 2017.
- [8] S. Ragi and E. K. P. Chong, "UAV path planning in a dynamic environment via partially observable Markov decision process," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 4, pp. 2397–2412, Oct. 2014.
- [9] M. Zhang and H. H. T. Liu, "Cooperative tracking a moving target using multiple fixed-wing UAVs," *J. Intell. Robot. Syst.*, vol. 81, nos. 3–4, p. 505, 2016.
- [10] T. Oliveira, A. P. Aguiar, and P. Encarnacao, "Moving path following for unmanned aerial vehicles with applications to single and multiple target tracking problems," *IEEE Trans. Robot.*, vol. 32, no. 5, pp. 1062–1078, Oct. 2016.
- [11] N. Farmani, L. Sun, and D. J. Pack, "A scalable multitarget tracking system for cooperative unmanned aerial vehicles," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 4, pp. 1947–1961, Aug. 2017.
- [12] X. Wang and L. Cai, "Navigation and control based on integral-uncertainty observer for unmanned jet aircraft," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1230–1249, Jun. 2017.
- [13] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960.
- [14] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Chelmsford, MA, USA: Courier Corporation, 2007.



- [15] E. B. Quist and R. W. Beard, "Radar odometry on fixed-wing small unmanned aircraft," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 1, pp. 396–410, Feb. 2016.
- [16] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.
- [17] L. Wang, Y. Li, H. Zhu, and L. Shen, "Target state estimation and prediction based standoff tracking of ground moving target using a fixed-wing UAV," in *Proc. IEEE Int. Conf. Control Automat.*, Jun. 2010, pp. 273–278.
- [18] J. T. Ambadan and Y. Tang, "Sigma-point Kalman filter data assimilation methods for strongly nonlinear systems," *J. Atmos. Sci.*, vol. 66, no. 2, pp. 261–285, Feb. 2009.
- [19] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.
- [20] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 910–927, May 2000.
- [21] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. Amer. Stat. Assoc.*, vol. 93, no. 443, pp. 1032–1044, Aug. 1998.
- [22] C. Hue, J.-P. Le Cadre, and P. Pérez, "Tracking multiple objects with particle filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 3, pp. 791–812, Jul. 2002.
- [23] J. H. Kotecha and P. M. Djuric, "Gaussian sum particle filtering for dynamic state space models," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 6, May 2001, pp. 3465–3468.
- [24] 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. Conf. Acoust., Speech, Signal Process.*, vol. 6, Apr. 2003, pp. VI-701.
- [25] Y. Huang, Y. Zhang, N. Li, and J. Chambers, "Robust student's t based nonlinear filter and smoother," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 5, pp. 2586–2596, Oct. 2016.
- [26] Y. Huang and Y. Zhang, "Robust student's t-based stochastic cubature filter for nonlinear systems with heavy-tailed process and measurement noises," *IEEE Access*, vol. 5, pp. 7964–7974, 2017.
- [27] B. Liao, Z. G. Zhang, and S. C. Chan, "A new robust Kalman filter-based subspace tracking algorithm in an impulsive noise environment," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 9, pp. 740–744, Sep. 2010.
- [28] S. C. Chan, B. Liao, and K. M. Tsui, "Bayesian Kalman filtering, regularization and compressed sampling," in *Proc. IEEE 54th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2011, pp. 1–4.
- [29] S. Zhang, C. Wang, S.-C. Chan, X. Wei, and C.-H. Ho, "New object detection, tracking, and recognition approaches for video surveillance over camera network," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2679–2691, May 2015.
- [30] S. Zhang, Z. Zhang, S. C. Chan, H. Wen, and X. Chen, "An automatic muscle fiber orientation tracking algorithm using Bayesian Kalman filter for ultrasound images," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 3510–3514.
- [31] K. K. Kottakki, M. Bhushan, and S. Bhartiya, "An improved Gaussian sum unscented Kalman filter," in *Proc. 3rd Int. Conf. Adv. Control Optim. Dyn. Syst.*, 2014, pp. 355–362.
- [32] M. Simandl and J. Duník, "Sigma point Gaussian sum filter design using square root unscented filters," *IFAC Proc. Volumes*, vol. 38, no. 1, pp. 1000–1005, 2005.
- [33] Z. Liu, S.-C. Chan, H.-C. Wu, and J. Wu, "Bayesian unscented Kalman filter for state estimation of nonlinear and non-Gaussian systems," in *Proc. 24th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2016, pp. 443–447.
- [34] K. H. Kim, J. G. Lee, and C. G. Park, "Adaptive two-stage Kalman filter in the presence of unknown random bias," *Int. J. Adapt. Control Signal Process.*, vol. 20, no. 7, pp. 305–319, 2006.
- [35] K.-H. Kim, G.-I. Jee, C.-G. Park, and J.-G. Lee, "The stability analysis of the adaptive fading extended Kalman filter using the innovation covariance," *Int. J. Control, Autom. Syst.*, vol. 7, no. 1, pp. 49–56, Feb. 2009.
- [36] K. H. Kim, J. G. Lee, and C. G. Park, "Adaptive two-stage extended Kalman filter for a fault-tolerant INS-GPS loosely coupled system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 1, pp. 125–137, Jan. 2009.
- [37] K.-H. Kim, G.-I. Jee, and J.-H. Song, "Carrier tracking loop using the adaptive two-stage Kalman filter for high dynamic situations," *Int. J. Control, Autom., Syst.*, vol. 6, no. 6, pp. 948–953, 2008.
- [38] S. Särkkä and J. Hartikainen, "Variational Bayesian adaptation of noise covariances in non-linear Kalman filtering," 2013, *arXiv:1302.0681*. [Online]. Available: <http://arxiv.org/abs/1302.0681>
- [39] J. L. Williams and P. S. Maybeck, "Cost-function-based hypothesis control techniques for multiple hypothesis tracking," *Math. Comput. Model.*, vol. 43, nos. 9–10, pp. 976–989, May 2006.
- [40] A. R. Runnalls, "Kullback-leibler approach to Gaussian mixture reduction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 3, pp. 989–999, Jul. 2007.
- [41] K. Zhang and J. T. Kwok, "Simplifying mixture models through function approximation," *IEEE Trans. Neural. Netw.*, vol. 21, no. 4, pp. 58–644, Apr. 2010.
- [42] Q. Xia, M. Rao, Y. Ying, and X. Shen, "Adaptive fading Kalman filter with an application," *Automatica*, vol. 30, no. 8, pp. 1333–1338, Aug. 1994.
- [43] C. Hide, T. Moore, and M. Smith, "Adaptive Kalman filtering for low-cost INS/GPS," *J. Navigat.*, vol. 56, no. 1, pp. 143–152, Jan. 2003.
- [44] L. Ozbek and F. A. Aliev, "Comments on adaptive fading Kalman filter with an application," *Automatica*, vol. 34, no. 12, pp. 1663–1664, 1998.
- [45] M. Efe and L. Ozbek, "Fading Kalman filter for manoeuvring target tracking," *J. Turkish Stat. Assoc.*, vol. 2, no. 3, pp. 193–206, 2012.
- [46] B. D. Anderson and J. B. Moore, *Optimal Filtering*. New York, NY, USA: Courier Corporation, 2012.
- [47] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Student J. Amer. Inst. Aeronaut. Astronaut.*, vol. 3, no. 8, pp. 1445–1450, Aug. 1965.
- [48] G. Kitagawa, "Non-Gaussian state-space modeling of nonstationary time series," *J. Amer. Stat. Assoc.*, vol. 82, no. 400, pp. 1032–1041, Dec. 1987.
- [49] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, no. 1, pp. 99–109, 1943.
- [50] S. S. Gleason and D. Gebre-Egziabher, *GNSS Applications and Methods*. Norwood, MA, USA: Artech House, 2009.



**ZHONG LIU** received the B.Eng. degree from the Huaihai Institute of Technology, Lianyungang, China, in 2010, the M.Eng. degree from the Zhongyuan University of Technology, Zhengzhou, China, in 2013, and the Ph.D. degree from the Department of Electrical and Electronics Engineering, The University of Hong Kong, Hong Kong, in 2018. He is currently a Research Fellow of the Department of Biomedical Engineering, Shenzhen University, Shenzhen, China. His research interests include signal processing, pattern recognition, and human-computer interaction.



**SHING-CHOW CHAN** (Member, IEEE) received the B.Sc.Eng. and Ph.D. degrees in electrical engineering from The University of Hong Kong, in 1986 and 1992, respectively. Since 1994, he has been with the Department of Electrical and Electronics Engineering, The University of Hong Kong, where he is currently a Professor and the Associate Head of the Department. His research interests include fast transform algorithms, filter design and realization, multirate and biomedical signal processing, communications and array signal processing, high-speed A/D converter architecture, bioinformatics, smart grids, and image-based rendering. He is currently a member of the Digital Signal Processing Technical Committee, IEEE Circuits and Systems Society. He was the Chair of the IEEE Hong Kong Chapter of Signal Processing, from 2000 to 2002, and an Organizing Committee Member of the IEEE ICASSP, in 2003, and the IEEE ICIP, in 2010. He served as an Associate Editor for *Digital Signal Processing*, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I: FUNDAMENTAL THEORY AND APPLICATIONS, from 2008 to 2009, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: EXPRESS BRIEFS, from 2012 to 2016. He is an Associate Editor of the *Journal of Signal Processing Systems*.