Huang Y, Zhang Y, Li N, Chambers JA.

Robust Student's t Based Nonlinear Filter and Smoother.

IEEE Transactions Aerospace and Electronic Systems 2016, 52(5), 2586-2596

DOI link to article:

http://dx.doi.org/10.1109/TAES.2016.150722

Date deposited:

08/02/2017

Robust Student's t based nonlinear filter and smoother

Yulong Huang, Yonggang Zhang, *Member, IEEE,* Ning Li, Jonathon Chambers *Fellow, IEEE*

***Abstract*—Novel Student's t based approaches for formulating a filter and smoother, which utilize heavy tailed process and measurement noise models, are found through approximations of the associated posterior probability density functions. Simulation results for manoeuvring target tracking illustrate that the proposed methods substantially outperform existing methods in terms of the root mean square error.**

***Index Terms*—State estimation, heavy tailed noise, Student's t based approximate filter, Student's t based approximate smoother, Student's t weighted integral, unscented transform**

I. Introduction

NONLINEAR filtering and smoothing have been widely used in many applications such as target tracking, control, signal processing and navigation [1], [2]. For general nonlinear systems, closed form solutions of posterior filtering and smoothing probability density functions (PDFs) are not available, thus optimal solutions normally don't exist and approximate approaches are necessary to design a suboptimal nonlinear filter or smoother [3]. Gaussian approximations to such PDFs are most common because their corresponding Gaussian approximate (GA) filter and smoother provide tradeoffs between computational complexity and estimation accuracy in many practical applications [4]–[7]. So far, several forms of GA filter and smoother have been developed using different rules, such as unscented Kalman filter (UKF) [8] and unscented Kalman smoother (UKS) [9] based on the unscented transform (UT); cubature Kalman filter (CKF) [5], [10] and cubature Kalman smoother [11], [12] utilizing a spherical radial cubature rule; embedded CKF based on embedded cubature rule [13] and interpolatory CKF based on interpolatory cubature rule [3]. These approaches realize a GA filter and smoother but are only suitable for nonlinear systems with Gaussian process and measurement noises. However, in some engineering applications, such as tracking agile targets with measurement outliers from unreliable sensors, their process and measurement noises are not Gaussian since they have heavy tails [14]. Thus, existing GA filters and smoothers may fail in some engineering applications with heavy tailed process and measurement noises.

Many linear approaches for realizing filters and smoothers have been derived based on the variational Bayesian approach to solve the filtering and smoothing problems of linear systems with heavy tailed measurement noises [15]–[17]. Robust nonlinear filters and smoothers can be designed for nonlinear systems with measurement outliers by using a combination of the multivariate Student's t distribution and the variational Bayesian approach [18]. Many Student's t based visual tracking methods have been proposed to deal with data outliers induced by varying object appearance, occlusions and changes in illumination [16], [19], [20]. However, these filters and smoothers are not suitable for the case of heavy tailed process noise since they are all based on the assumption of well behaved process noise [14]. To address the filtering problem of linear systems with heavy tailed process and measurement noises, a linear Student's t based filter was proposed by modelling both the process and measurement noises as Student's t distributions and approximating the posterior filtering PDF as a Student's t distribution [14]. Although this linear Student's t based filter can be used to achieve the filtering estimate for nonlinear systems with heavy tailed process and measurement noises based on the first order linearisation approach, it shows poor filtering performance since enormous truncation errors can be induced by the linearisation. Moreover, up to the present, Student's t based approaches to approximate a filter and smoother for nonlinear systems with heavy tailed process and measurement noises do not exist.

In this work, both a robust Student's t based nonlinear filter and smoother are derived by providing Student's t approximations to posterior filtering and smoothing PDFs. These can be deemed as a generalization and extension of the linear Student's t filter. The moment matching approach is used to constrain the increase of the degrees of freedom (dof) parameters of posterior filtering and smoothing PDFs, and the UT is used to compute the Student's t weighted integrals involved in the proposed Student's t based approximate filter and smoother. It is shown that existing approaches to linear Student's t based filter [14] and GA filter [5] are special cases of the proposed Student's t based filter, and the existing GA smoother [21] is a special case of the proposed Student's t based smoother. The performance of the proposed methods is tested using a manoeuvring target tracking scenario. Simulation results show that the proposed methods outperform existing methods for the case of heavy tailed process and measurement noises.

This work was supported by the National Natural Science Foundation of China under Grant Nos. 61201409 and 61371173 and the Fundamental Research Funds for the Central Universities of Harbin Engineering University No. HEUCFQ20150407 and the Engineering and Physical Sciences Research Council (EPSRC) of the UK grant no. EP/K014307/1.

Y. L. Huang, Y. G. Zhang and N. Li are with the Department of Automation, Harbin Engineering University, Harbin 150001, China (e-mail: heuedu@163.com; zhangyg@hrbeu.edu.cn; ningli@hrbeu.edu.cn). Y. G. Zhang is the corresponding author of this paper.

J. Chambers is with the School of Electrical and Electronic Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK (e-mail: Jonathon.Chambers@newcastle.ac.uk).

II. PROBLEM FORMULATION

Consider the following discrete-time nonlinear stochastic system as shown by the state-space model

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \qquad \text{(process equation)} \quad (1)$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \qquad \text{(measurement equation)}, \quad (2)$$

where k is the discrete time index, $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector, $\mathbf{z}_k \in \mathbb{R}^m$ is the measurement vector, $\mathbf{w}_k \in \mathbb{R}^n$ is the process noise vector, and $\mathbf{v}_k \in \mathbb{R}^m$ is the measurement noise vector. Process and measurement noise distributions are assumed to have heavy tails, and they are modelled as Student's t distributions as follows

$$p(\mathbf{w}_k) = \mathrm{St}(\mathbf{w}_k; \mathbf{0}, \mathbf{Q}_k, v_1) \quad (3)$$

$$p(\mathbf{v}_k) = \mathrm{St}(\mathbf{v}_k; \mathbf{0}, \mathbf{R}_k, v_2), \quad (4)$$

where $\mathrm{St}(\mathbf{x}; \mu, \mathbf{\Sigma}, v)$ denotes the Student's t PDF with mean vector μ, scale matrix $\mathbf{\Sigma}$, and dof parameter v. The initial state vector $\mathbf{x}_0$ is assumed to have a Student's t distribution with mean vector $\hat{\mathbf{x}}_{0|0}$, scale matrix $\mathbf{P}_{0|0}$, and dof parameter v_3, i.e.,

$$p(\mathbf{x}_0) = \mathrm{St}(\mathbf{x}_0; \hat{\mathbf{x}}_{0|0}, \mathbf{P}_{0|0}, v_3). \quad (5)$$

Moreover, $\mathbf{x}_0$, $\mathbf{w}_k$ and $\mathbf{v}_k$ are assumed to be mutually uncorrelated in this work.

Our aim is to provide Student's t based approaches to approximate a filter and smoother for the nonlinear systems formulated in (1)-(5). That is to say, we need to find Student's t based approximations to the filtering PDF $p(\mathbf{x}_k|\mathbf{Z}_k)$ and smoothing PDF $p(\mathbf{x}_k|\mathbf{Z}_N)$, i.e.,

$$p(\mathbf{x}_k|\mathbf{Z}_k) = \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}, v_3) \quad (6)$$

$$p(\mathbf{x}_k|\mathbf{Z}_N) = \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|N}, \mathbf{P}_{k|N}, v_3), \quad (7)$$

where $\mathbf{Z}_N = \{\mathbf{z}_j\}_{j=1}^N$ is the set of all N measurement vectors, $\hat{\mathbf{x}}_{k|k}$ is the filtering estimate vector, $\mathbf{P}_{k|k}$ is the scale matrix of the filtering PDF, $\hat{\mathbf{x}}_{k|N}$ is the smoothing estimate vector, and $\mathbf{P}_{k|N}$ is the scale matrix of the smoothing PDF which can be approximated as

$$\hat{\mathbf{x}}_{k|k} = E[\mathbf{x}_k|\mathbf{Z}_k] = \int \mathbf{x}_k p(\mathbf{x}_k|\mathbf{Z}_k) d\mathbf{x}_k \quad (8)$$

$$\mathbf{P}_{k|k} = \frac{v_3 - 2}{v_3} E[\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T|\mathbf{Z}_k] = \frac{v_3 - 2}{v_3} \int \tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T \times p(\mathbf{x}_k|\mathbf{Z}_k) d\mathbf{x}_k \quad (9)$$

$$\hat{\mathbf{x}}_{k|N} = E[\mathbf{x}_k|\mathbf{Z}_N] = \int \mathbf{x}_k p(\mathbf{x}_k|\mathbf{Z}_N) d\mathbf{x}_k \quad (10)$$

$$\mathbf{P}_{k|N} = \frac{v_3 - 2}{v_3} E[\tilde{\mathbf{x}}_{k|N}\tilde{\mathbf{x}}_{k|N}^T|\mathbf{Z}_N] = \frac{v_3 - 2}{v_3} \int \tilde{\mathbf{x}}_{k|N}\tilde{\mathbf{x}}_{k|N}^T \times p(\mathbf{x}_k|\mathbf{Z}_N) d\mathbf{x}_k, \quad (11)$$

where $\tilde{\mathbf{x}}_{k|k} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}$ and $\tilde{\mathbf{x}}_{k|N} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|N}$ denote the estimate error and smoothing error vectors respectively, and equations (9) and (11) are obtained from the relationship between scale matrix and the covariance matrix [14]. (For a Student's t random vector $\mathbf{x}$ with PDF $p(\mathbf{x}) = \mathrm{St}(\mathbf{x}; \mu, \mathbf{\Sigma}, v)$, its covariance matrix is $\frac{v}{v-2}\mathbf{\Sigma}$, i.e., $E[(\mathbf{x}-\mu)(\mathbf{x}-\mu)^T] = \frac{v}{v-2}\mathbf{\Sigma}$. Conversely, $\mathbf{\Sigma} = \frac{v-2}{v}E[(\mathbf{x}-\mu)(\mathbf{x}-\mu)^T]$.)

To approximate posterior filtering and smoothing PDFs as Gaussian, the joint PDFs $p(\mathbf{x}_k, \mathbf{z}_k|\ \mathbf{Z}_{k-1})$ and $p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k)$ are assumed, respectively, to be Gaussian in the standard GA filter and smoother [5], [21]. However, in the case that process and measurement noise distributions have heavy tails, such assumptions are unreasonable since these joint PDFs are not Gaussian any more.

In order to approximate the posterior filtering and smoothing PDFs as Student's t, the joint PDFs $p(\mathbf{x}_k, \mathbf{z}_k|\ \mathbf{Z}_{k-1})$ and $p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k)$ need to be assumed Student's t. These Student's t assumptions hold for linear systems due to the affine property of the Student's t random vector, and they are reasonable for applications with mild nonlinearity, such as target tracking as will be shown in our later simulations.

Assumption 1: The jointly predicted PDF $p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{Z}_{k-1})$ of the state and measurement vectors is Student's t, i.e.,

$$p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{Z}_{k-1}) = \mathrm{St}(\begin{bmatrix}\mathbf{x}_k \\ \mathbf{z}_k\end{bmatrix}; \begin{bmatrix}\hat{\mathbf{x}}_{k|k-1} \\ \hat{\mathbf{z}}_{k|k-1}\end{bmatrix}, \begin{bmatrix}\mathbf{P}_{k|k-1} & \mathbf{P}_{k|k-1}^{xz} \\ (\mathbf{P}_{k|k-1}^{xz})^T & \mathbf{P}_{k|k-1}^{zz}\end{bmatrix}, v_3), \quad (12)$$

where $\hat{\mathbf{x}}_{k|k-1}$, $\mathbf{P}_{k|k-1}$, $\hat{\mathbf{z}}_{k|k-1}$, $\mathbf{P}_{k|k-1}^{zz}$ and $\mathbf{P}_{k|k-1}^{xz}$ can be computed as

$$\hat{\mathbf{x}}_{k|k-1} = E[\mathbf{x}_k|\mathbf{Z}_{k-1}] = \int \mathbf{x}_k p(\mathbf{x}_k|\mathbf{Z}_{k-1}) d\mathbf{x}_k \quad (13)$$

$$\mathbf{P}_{k|k-1} = \frac{v_3 - 2}{v_3} E[\tilde{\mathbf{x}}_{k|k-1}\tilde{\mathbf{x}}_{k|k-1}^T|\mathbf{Z}_{k-1}] = \frac{v_3 - 2}{v_3} \int \tilde{\mathbf{x}}_{k|k-1} \times \tilde{\mathbf{x}}_{k|k-1}^T p(\mathbf{x}_k|\mathbf{Z}_{k-1}) d\mathbf{x}_k \quad (14)$$

$$\hat{\mathbf{z}}_{k|k-1} = E[\mathbf{z}_k|\mathbf{Z}_{k-1}] = \int \mathbf{z}_k p(\mathbf{z}_k|\mathbf{Z}_{k-1}) d\mathbf{z}_k \quad (15)$$

$$\mathbf{P}_{k|k-1}^{zz} = \frac{v_3 - 2}{v_3} E[\tilde{\mathbf{z}}_{k|k-1}\tilde{\mathbf{z}}_{k|k-1}^T|\mathbf{Z}_{k-1}] = \frac{v_3 - 2}{v_3} \int \tilde{\mathbf{z}}_{k|k-1} \times \tilde{\mathbf{z}}_{k|k-1}^T p(\mathbf{z}_k|\mathbf{Z}_{k-1}) d\mathbf{z}_k \quad (16)$$

$$\mathbf{P}_{k|k-1}^{xz} = \frac{v_3 - 2}{v_3} E[\tilde{\mathbf{x}}_{k|k-1}\tilde{\mathbf{z}}_{k|k-1}^T|\mathbf{Z}_{k-1}] = \frac{v_3 - 2}{v_3} \int\int \tilde{\mathbf{x}}_{k|k-1}\tilde{\mathbf{z}}_{k|k-1}^T p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{Z}_{k-1}) d\mathbf{x}_k d\mathbf{z}_k, \quad (17)$$

where $\tilde{\mathbf{x}}_{k|k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}$ denotes the predicted error vectors, and $\tilde{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}$ denotes the innovation vector, and equations (14) and (16)-(17) are obtained from the relationship between the scale matrix and covariance matrix [14].

Assumption 2: The joint PDF $p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k)$ of the current state $\mathbf{x}_k$ and one-step ahead state $\mathbf{x}_{k+1}$ vectors is Student's t, i.e.,

$$p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k) = \mathrm{St}(\begin{bmatrix}\mathbf{x}_k \\ \mathbf{x}_{k+1}\end{bmatrix}; \begin{bmatrix}\hat{\mathbf{x}}_{k|k} \\ \hat{\mathbf{x}}_{k+1|k}\end{bmatrix}, \begin{bmatrix}\mathbf{P}_{k|k} & \mathbf{P}_{k,k+1|k} \\ \mathbf{P}_{k,k+1|k}^T & \mathbf{P}_{k+1|k}\end{bmatrix}, v_3), \quad (18)$$

where the scale matrix $\mathbf{P}_{k,k+1|k}$ can be computed as

$$\mathbf{P}_{k,k+1|k} = \frac{v_3 - 2}{v_3} E[\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k+1|k}^T|\mathbf{Z}_k] = \frac{v_3 - 2}{v_3} \int\int \tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k+1|k}^T p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k) d\mathbf{x}_k d\mathbf{x}_{k+1}. \quad (19)$$

Based on Assumptions 1-2, the novel Student's t based approximate filter and smoother for nonlinear systems with heavy tailed process and measurement noises can be derived.

III. A Student's t based approximate filter

A. Time Update

In the time update, the one-step predicted PDF $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ of the state is computed as follows [5]

$$p(\mathbf{x}_k|\mathbf{Z}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1})d\mathbf{x}_{k-1}. \tag{20}$$

Using (1), (3) and (6) in (20), we obtain

$$p(\mathbf{x}_k|\mathbf{Z}_{k-1}) = \int \mathrm{St}(\mathbf{x}_k; \mathbf{f}_{k-1}(\mathbf{x}_{k-1}), \mathbf{Q}_{k-1}, v_1) \times \mathrm{St}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}, v_3)d\mathbf{x}_{k-1} \tag{21}$$

It can be seen from (21) that $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ is not a Student's t PDF due to the nonlinear propagation $\mathbf{f}_{k-1}(\cdot)$. Thus, $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ needs to be approximated as a Student's t PDF based on Assumption 1

$$p(\mathbf{x}_k|\mathbf{Z}_{k-1}) = \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3). \tag{22}$$

Using (21) in (13)-(14), $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ can be computed as follows

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \int \mathbf{x}_k p(\mathbf{x}_k|\mathbf{Z}_{k-1})d\mathbf{x}_k \\ &= \int [\int \mathbf{x}_k \mathrm{St}(\mathbf{x}_k; \mathbf{f}_{k-1}(\mathbf{x}_{k-1}), \mathbf{Q}_{k-1}, v_1)d\mathbf{x}_k] \times \\ &\quad \mathrm{St}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}, v_3)d\mathbf{x}_{k-1} \\ &= \int \mathbf{f}_{k-1}(\mathbf{x}_{k-1})\mathrm{St}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}, v_3) \\ &\quad d\mathbf{x}_{k-1}\end{aligned} \tag{23}$$

$$\begin{aligned}\mathbf{P}_{k|k-1} &= \frac{v_3-2}{v_3}\int \mathbf{x}_k\mathbf{x}_k^T p(\mathbf{x}_k|\mathbf{Z}_{k-1})d\mathbf{x}_k - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k-1} \times \\ &\quad \hat{\mathbf{x}}_{k|k-1}^T \\ &= \frac{v_3-2}{v_3}\int [\int \mathbf{x}_k\mathbf{x}_k^T \mathrm{St}(\mathbf{x}_k; \mathbf{f}_{k-1}(\mathbf{x}_{k-1}), \mathbf{Q}_{k-1}, v_1) \\ &\quad d\mathbf{x}_k]\mathrm{St}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}, v_3)d\mathbf{x}_{k-1} - \\ &\quad \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}_{k|k-1}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{f}_{k-1}(\mathbf{x}_{k-1})\mathbf{f}_{k-1}^T(\mathbf{x}_{k-1}) \times \\ &\quad \mathrm{St}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}, v_3)d\mathbf{x}_{k-1} - \\ &\quad \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}_{k|k-1}^T + \frac{v_1(v_3-2)}{(v_1-2)v_3}\mathbf{Q}_{k-1}.\end{aligned} \tag{24}$$

The time update consists of (22)-(24), where the one-step predicted PDF $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ of the state is approximated as Student's t. To approximate the filtering PDF $p(\mathbf{x}_k|\mathbf{Z}_k)$ as Student's t, the measurement update will be derived as follows.

B. Measurement Update

In the measurement update, firstly, the likelihood PDF $p(\mathbf{z}_k|\mathbf{Z}_{k-1})$ is computed as follows

$$p(\mathbf{z}_k|\mathbf{Z}_{k-1}) = \int p(\mathbf{z}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Z}_{k-1})d\mathbf{x}_k. \tag{25}$$

Using (2), (4) and (22) in (25), we obtain

$$\begin{aligned}p(\mathbf{z}_k|\mathbf{Z}_{k-1}) &= \int \mathrm{St}(\mathbf{z}_k; \mathbf{h}_k(\mathbf{x}_k), \mathbf{R}_k, v_2) \times \\ &\quad \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3)d\mathbf{x}_k.\end{aligned} \tag{26}$$

It can be seen from (26) that $p(\mathbf{z}_k|\mathbf{Z}_{k-1})$ is not a Student's t PDF due to the nonlinear propagation $\mathbf{h}_k(\cdot)$. Thus, $p(\mathbf{z}_k|\mathbf{Z}_{k-1})$ is also approximated as a Student's t PDF based on Assumption 1

$$p(\mathbf{z}_k|\mathbf{Z}_{k-1}) = \mathrm{St}(\mathbf{z}_k; \hat{\mathbf{z}}_{k|k-1}, \mathbf{P}_{k|k-1}^{zz}, v_3). \tag{27}$$

Using (26) in (15)-(16), $\hat{\mathbf{z}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}^{zz}$ can be computed as

$$\begin{aligned}\hat{\mathbf{z}}_{k|k-1} &= \int \mathbf{z}_k p(\mathbf{z}_k|\mathbf{Z}_{k-1})d\mathbf{z}_k \\ &= \int [\int \mathbf{z}_k \mathrm{St}(\mathbf{z}_k; \mathbf{h}_k(\mathbf{x}_k), \mathbf{R}_k, v_2)d\mathbf{z}_k] \times \\ &\quad \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3)d\mathbf{x}_k \\ &= \int \mathbf{h}_k(\mathbf{x}_k)\mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3)d\mathbf{x}_k\end{aligned} \tag{28}$$

$$\begin{aligned}\mathbf{P}_{k|k-1}^{zz} &= \frac{v_3-2}{v_3}\int \mathbf{z}_k\mathbf{z}_k^T p(\mathbf{z}_k|\mathbf{Z}_{k-1})d\mathbf{z}_k - \\ &\quad \frac{v_3-2}{v_3}\hat{\mathbf{z}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T \\ &= \frac{v_3-2}{v_3}\int [\int \mathbf{z}_k\mathbf{z}_k^T \mathrm{St}(\mathbf{z}_k; \mathbf{h}_k(\mathbf{x}_k), \mathbf{R}_k, v_2)d\mathbf{z}_k] \times \\ &\quad \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3)d\mathbf{x}_k - \\ &\quad \frac{v_3-2}{v_3}\hat{\mathbf{z}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{h}_k(\mathbf{x}_k)\mathbf{h}_k^T(\mathbf{x}_k)\mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3) \\ &\quad d\mathbf{x}_k - \frac{v_3-2}{v_3}\hat{\mathbf{z}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T + \frac{v_2(v_3-2)}{(v_2-2)v_3}\mathbf{R}_k.\end{aligned} \tag{29}$$

It can be seen from (12) that, to approximate $p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{Z}_{k-1})$ as a Student's t PDF, we need to compute $\mathbf{P}_{k|k-1}^{xz}$ using (17) as follows

$$\begin{aligned}\mathbf{P}_{k|k-1}^{xz} &= \frac{v_3-2}{v_3}\int\int \mathbf{x}_k\mathbf{z}_k^T p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{Z}_{k-1})d\mathbf{z}_k d\mathbf{x}_k - \\ &\quad \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{x}_k[\int \mathbf{z}_k^T p(\mathbf{z}_k|\mathbf{x}_k)d\mathbf{z}_k]p(\mathbf{x}_k|\mathbf{Z}_{k-1})d\mathbf{x}_k - \\ &\quad \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{x}_k[\int \mathbf{z}_k^T \mathrm{St}(\mathbf{z}_k; \mathbf{h}_k(\mathbf{x}_k), \mathbf{R}_k, v_2)d\mathbf{z}_k] \times \\ &\quad \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3)d\mathbf{x}_k - \\ &\quad \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{x}_k\mathbf{h}_k^T(\mathbf{x}_k)\mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}, v_3)d\mathbf{x}_k \\ &\quad - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{z}}_{k|k-1}^T.\end{aligned} \tag{30}$$

Secondly, according to Bayes rule and using (12) and (27), the posterior filtering PDF $p(\mathbf{x}_k|\mathbf{Z}_k)$ can be updated as a Student's t PDF, i.e., [14]

$$p(\mathbf{x}_k|\mathbf{Z}_k) = \frac{p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{Z}_{k-1})}{p(\mathbf{z}_k|\mathbf{Z}_{k-1})} = \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}^{'}_{k|k}, \mathbf{P}^{'}_{k|k}, v^{'}_3), \quad (31)$$

where $\hat{\mathbf{x}}^{'}_{k|k}$, $\mathbf{P}^{'}_{k|k}$ and $v^{'}_3$ are given by

$$v^{'}_3 = v_3 + m \quad (32)$$

$$\hat{\mathbf{x}}^{'}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}) \quad (33)$$

$$\mathbf{P}^{'}_{k|k} = \frac{v_3 + \Delta^2_k}{v_3 + m}(\mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}^{zz}_{k|k-1} \mathbf{K}^T_k) \quad (34)$$

$$\mathbf{K}_k = \mathbf{P}^{xz}_{k|k-1}(\mathbf{P}^{zz}_{k|k-1})^{-1} \quad (35)$$

$$\Delta_k = \sqrt{(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^T(\mathbf{P}^{zz}_{k|k-1})^{-1}(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})}. \quad (36)$$

It is seen from (32) that the dof parameter $v^{'}_3$ of $p(\mathbf{x}_k|\mathbf{Z}_k)$ will tend to infinity as the time increases so that $p(\mathbf{x}_k|\mathbf{Z}_k)$ will converge to a Gaussian PDF. In other words, $p(\mathbf{x}_k|\mathbf{Z}_k)$ will lose the heavy tailed properties that must be retained. Similar to [14], the moment matching approach is used to solve this problem. Here, we only need to match the first two moments, i.e.,

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}^{'}_{k|k} \qquad \frac{v_3}{v_3 - 2}\mathbf{P}_{k|k} = \frac{v^{'}_3}{v^{'}_3 - 2}\mathbf{P}^{'}_{k|k}. \quad (37)$$

According to (37), $\mathbf{P}_{k|k}$ can be approximated as follows

$$\mathbf{P}_{k|k} = \frac{(v_3 - 2)v^{'}_3}{v_3(v^{'}_3 - 2)}\mathbf{P}^{'}_{k|k}. \quad (38)$$

The proposed Student's t based approximate filter operates by combining the analytical computations in (32)-(38) with the Student's t weighted integrals in (23)-(24) and (28)-(30).

C. Comparisons with existing linear Student's t based filter and GA filter

In [14], a Student's t based filter is designed for linear systems with heavy tailed process and measurement noises. It is interesting that the Student's t based filter [14] is a special case of the proposed filter when $v_1 = v_2 = v_3 = v$, which is confirmed as follows. For a linear system, we let

$$\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) = \mathbf{F}_{k-1}\mathbf{x}_{k-1} \quad \mathbf{h}_k(\mathbf{x}_k) = \mathbf{H}_k\mathbf{x}_k. \quad (39)$$

Substituting (39) in (23)-(24) and (28)-(30) and using $v_1 = v_2 = v_3 = v$, we obtain

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} \quad (40)$$

$$\mathbf{P}_{k|k-1} = \frac{v-2}{v}\mathbf{F}_{k-1}(\hat{\mathbf{x}}_{k-1|k-1}\hat{\mathbf{x}}^T_{k-1|k-1} + \frac{v}{v-2}\mathbf{P}_{k-1|k-1}) \mathbf{F}^T_{k-1} - \frac{v-2}{v}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}^T_{k|k-1} + \mathbf{Q}_{k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}^T_{k-1} + \mathbf{Q}_{k-1} \quad (41)$$

$$\hat{\mathbf{z}}_{k|k-1} = \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1} \quad (42)$$

$$\mathbf{P}^{zz}_{k|k-1} = \frac{v-2}{v}\mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}^T_{k|k-1} + \frac{v}{v-2}\mathbf{P}_{k|k-1})\mathbf{H}^T_k - \frac{v-2}{v}\hat{\mathbf{z}}_{k|k-1}\hat{\mathbf{z}}^T_{k|k-1} + \mathbf{R}_k = \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}^T_k + \mathbf{R}_k \quad (43)$$

$$\mathbf{P}^{xz}_{k|k-1} = \frac{v-2}{v}(\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}^T_{k|k-1} + \frac{v}{v-2}\mathbf{P}_{k|k-1})\mathbf{H}^T_k - \frac{v-2}{v}\hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{z}}^T_{k|k-1} = \mathbf{P}_{k|k-1}\mathbf{H}^T_k. \quad (44)$$

Using (40)-(44) in (32)-(36), the existing linear Student's t based filter can be obtained. Thus, the existing linear Student's t based filter is a special case of the proposed filter when $v_1 = v_2 = v_3 = v$.

Moreover, the proposed filter converges to the existing GA filter [5] when $v_1 = v_2 = v_3 = v \to +\infty$, which is verified as follows. If $v_1 = v_2 = v_3 = v \to +\infty$, we have

$$\frac{v_3 - 2}{v_3} \to 1 \qquad \frac{v_3 + \Delta^2_k}{v_3 + m} \to 1 \quad (45)$$

$$\mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}, v_3) \to N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}) \quad (46)$$

$$\mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k}, \mathbf{P}_{k+1|k}, v_3) \to N(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|k}, \mathbf{P}_{k+1|k}), \quad (47)$$

where $N(\mathbf{x}; \mu, \mathbf{\Sigma})$ denotes the Gaussian PDF with mean vector μ and covariance matrix $\mathbf{\Sigma}$.

Substituting (45)-(47) in (23)-(24) and (28)-(30), we obtain

$$\lim_{v\to+\infty} \hat{\mathbf{x}}_{k|k-1} = \int \mathbf{f}_{k-1}(\mathbf{x}_{k-1})N(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})d\mathbf{x}_{k-1} \quad (48)$$

$$\lim_{v\to+\infty} \mathbf{P}_{k|k-1} = \int \mathbf{f}_{k-1}(\mathbf{x}_{k-1})\mathbf{f}^T_{k-1}(\mathbf{x}_{k-1})N(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})d\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{x}}^T_{k|k-1} + \mathbf{Q}_{k-1} \quad (49)$$

$$\lim_{v\to+\infty} \hat{\mathbf{z}}_{k|k-1} = \int \mathbf{h}_k(\mathbf{x}_k)N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})d\mathbf{x}_k \quad (50)$$

$$\lim_{v\to+\infty} \mathbf{P}^{zz}_{k|k-1} = \int \mathbf{h}_k(\mathbf{x}_k)\mathbf{h}^T_k(\mathbf{x}_k)N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) d\mathbf{x}_k - \hat{\mathbf{z}}_{k|k-1}\hat{\mathbf{z}}^T_{k|k-1} + \mathbf{R}_k \quad (51)$$

$$\lim_{v\to+\infty} \mathbf{P}^{xz}_{k|k-1} = \int \mathbf{x}_k\mathbf{h}^T_k(\mathbf{x}_k)N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})d\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\hat{\mathbf{z}}^T_{k|k-1}. \quad (52)$$

Substituting (32) and (34) in (38) and using $v_1 = v_2 = v_3 = v \to +\infty$, we obtain

$$\lim_{v\to+\infty} \mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{P}^{zz}_{k|k-1}\mathbf{K}^T_k. \quad (53)$$

It is seen from (33), (35), (37) and (48)-(53) that the state estimate vector $\hat{\mathbf{x}}_{k|k}$ and scale matrix $\mathbf{P}_{k|k}$ of the proposed

Student's t based approximate filter will converge to the state estimate vector $\hat{\mathbf{x}}_{k|k}^{GA}$ and corresponding estimate error covariance matrix $\mathbf{P}_{k|k}^{GA}$ of an existing GA filter as the dof parameters $v_1 = v_2 = v_3 = v \to +\infty$. Considering that the Student's t PDF converges to a Gaussian PDF as the dof parameters tends to infinity results in

$$\lim_{v\to+\infty} p(\mathbf{x}_k|\mathbf{Z}_k) = \lim_{v\to+\infty} \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}, v) = N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{GA}, \mathbf{P}_{k|k}^{GA}). \quad (54)$$

Equation (54) implies that the existing GA filter can be deemed as a special case of the proposed Student's t based approximation filter when $v_1 = v_2 = v_3 = v \to +\infty$.

IV. A Student's t based approximate smoother

According to Bayes theorem, the smoothing PDF $p(\mathbf{x}_k|\mathbf{Z}_N)$ can be computed as

$$p(\mathbf{x}_k|\mathbf{Z}_N) = \int p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_N) d\mathbf{x}_{k+1} = \int p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{Z}_N) p(\mathbf{x}_{k+1}|\mathbf{Z}_N) d\mathbf{x}_{k+1}. \quad (55)$$

The Markov property of the model in (1)-(2) implies that given the knowledge of $\mathbf{x}_{k+1}$ and $\mathbf{Z}_k$, the state vector $\mathbf{x}_k$ is independent of the future measurements $\{\mathbf{z}_j\}_{j=k+1}^N$, i.e., [21]

$$p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{Z}_N) = p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{Z}_k). \quad (56)$$

Substituting (56) in (55), $p(\mathbf{x}_k|\mathbf{Z}_N)$ can be written as

$$p(\mathbf{x}_k|\mathbf{Z}_N) = \int p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{Z}_k) p(\mathbf{x}_{k+1}|\mathbf{Z}_N) d\mathbf{x}_{k+1}. \quad (57)$$

It can be seen from (57) that, to approximate $p(\mathbf{x}_k|\mathbf{Z}_N)$ as a Student's t PDF, it is necessary to compute $p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{Z}_k)$ based on Assumption 2. Firstly, to approximate $p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k)$ as a Student's t PDF, we need to compute $\mathbf{P}_{k,k+1|k}$ using (19) as follows

$$\begin{aligned}\mathbf{P}_{k,k+1|k} &= \frac{v_3-2}{v_3} E[\mathbf{x}_k\mathbf{x}_{k+1}^T|\mathbf{Z}_k] - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k+1|k}^T \\ &= \frac{v_3-2}{v_3}\int\int \mathbf{x}_k\mathbf{x}_{k+1}^T p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k) d\mathbf{x}_{k+1} d\mathbf{x}_k - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k+1|k}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{x}_k [\int \mathbf{x}_{k+1}^T p(\mathbf{x}_{k+1}|\mathbf{x}_k) d\mathbf{x}_{k+1}] \times p(\mathbf{x}_k|\mathbf{Z}_k) d\mathbf{x}_k - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k+1|k}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{x}_k [\int \mathbf{x}_{k+1}^T \mathrm{St}(\mathbf{x}_{k+1}; \mathbf{f}_k(\mathbf{x}_k), \mathbf{Q}_k, v_1) d\mathbf{x}_{k+1}] \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}, v_3) d\mathbf{x}_k - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k+1|k}^T \\ &= \frac{v_3-2}{v_3}\int \mathbf{x}_k \mathbf{f}_k^T(\mathbf{x}_k) \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}, v_3) d\mathbf{x}_k - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k+1|k}^T. \end{aligned} \quad (58)$$

Secondly, according to Bayes rule and using (18) and (22), $p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{Z}_k)$ can be updated as a Student's t PDF, i.e., [14]

$$p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{Z}_k) = \frac{p(\mathbf{x}_k, \mathbf{x}_{k+1}|\mathbf{Z}_k)}{p(\mathbf{x}_{k+1}|\mathbf{Z}_k)} = \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k+1,k}, \mathbf{P}_{k|k+1,k}, v_3^{''}), \quad (59)$$

where $\hat{\mathbf{x}}_{k|k+1,k}$, $\mathbf{P}_{k|k+1,k}$ and $v_3^{''}$ are given by

$$v_3^{''} = v_3 + n \quad (60)$$

$$\hat{\mathbf{x}}_{k|k+1,k} = \hat{\mathbf{x}}_{k|k} + \mathbf{A}_k(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k}) \quad (61)$$

$$\mathbf{P}_{k|k+1,k} = \frac{v_3 + \Lambda_k^2}{v_3 + n}(\mathbf{P}_{k|k} - \mathbf{A}_k\mathbf{P}_{k+1|k}\mathbf{A}_k^T) \quad (62)$$

$$\mathbf{A}_k = \mathbf{P}_{k,k+1|k}\mathbf{P}_{k+1|k}^{-1} \quad (63)$$

$$\Lambda_k = \sqrt{(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T\mathbf{P}_{k+1|k}^{-1}(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})}. \quad (64)$$

Substituting (7) and (59) in (57), we obtain

$$p(\mathbf{x}_k|\mathbf{Z}_N) = \int \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k+1,k}, \mathbf{P}_{k|k+1,k}, v_3^{''}) \times \mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \mathbf{P}_{k+1|N}, v_3) d\mathbf{x}_{k+1}. \quad (65)$$

It is seen from (65) that the smoothing PDF $p(\mathbf{x}_k|\mathbf{Z}_N)$ is not a Student's t distribution since $\mathbf{P}_{k|k+1,k}$ is a quadratic function of $\mathbf{x}_{k+1}$. Here, $p(\mathbf{x}_k|\mathbf{Z}_N)$ is approximated as a Student's t PDF in (7). By substituting (65) in (10)-(11) and using (61)-(64), $\hat{\mathbf{x}}_{k|N}$ and $\mathbf{P}_{k|N}$ are computed as

$$\begin{aligned}\hat{\mathbf{x}}_{k|N} &= \int \mathbf{x}_k p(\mathbf{x}_k|\mathbf{Z}_N) d\mathbf{x}_k \\ &= \int\int \mathbf{x}_k \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k+1,k}, \mathbf{P}_{k|k+1,k}, v_3^{''}) d\mathbf{x}_k \times \mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \mathbf{P}_{k+1|N}, v_3) d\mathbf{x}_{k+1} \\ &= \int \hat{\mathbf{x}}_{k|k+1,k} \mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \mathbf{P}_{k+1|N}, v_3) d\mathbf{x}_{k+1} \\ &= \int [\hat{\mathbf{x}}_{k|k} + \mathbf{A}_k(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})] \times \mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \mathbf{P}_{k+1|N}, v_3) d\mathbf{x}_{k+1} \\ &= \hat{\mathbf{x}}_{k|k} + \mathbf{A}_k(\hat{\mathbf{x}}_{k+1|N} - \hat{\mathbf{x}}_{k+1|k}) \end{aligned} \quad (66)$$

$$\mathbf{P}_{k|N} = \frac{v_3-2}{v_3} E[\mathbf{x}_k\mathbf{x}_k^T|\mathbf{Z}_N] - \frac{v_3-2}{v_3}\hat{\mathbf{x}}_{k|N}\hat{\mathbf{x}}_{k|N}^T, \quad (67)$$

where $E[\mathbf{x}_k\mathbf{x}_k^T|\mathbf{Z}_N]$ can be computed as

$$\begin{aligned} E[\mathbf{x}_k\mathbf{x}_k^T|\mathbf{Z}_N] &= \int\int \mathbf{x}_k\mathbf{x}_k^T \mathrm{St}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k+1,k}, \mathbf{P}_{k|k+1,k}, v_3^{''}) d\mathbf{x}_k \times \mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \mathbf{P}_{k+1|N}, v_3) d\mathbf{x}_{k+1} \\ &= \int [\hat{\mathbf{x}}_{k|k+1,k}\hat{\mathbf{x}}_{k|k+1,k}^T + \frac{v_3^{''}}{v_3^{''}-2}\mathbf{P}_{k|k+1,k}] \times \mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \mathbf{P}_{k+1|N}, v_3) d\mathbf{x}_{k+1} \\ &= \mathbf{\Gamma} + \frac{v_3^{''}}{v_3^{''}-2}\mathbf{\Omega}. \end{aligned} \quad (68)$$

According to the definition of $\mathbf{\Gamma}$ in (68), we obtain

$$\begin{aligned}\mathbf{\Gamma} &= \int [\mathbf{A}_k(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|N}) + \hat{\mathbf{x}}_{k|N}][\mathbf{A}_k(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|N}) + \hat{\mathbf{x}}_{k|N}]^T \mathrm{St}(\mathbf{x}_{k+1}; \hat{\mathbf{x}}_{k+1|N}, \mathbf{P}_{k+1|N}, v_3) d\mathbf{x}_{k+1} \\ &= \hat{\mathbf{x}}_{k|N}\hat{\mathbf{x}}_{k|N}^T + \frac{v_3}{v_3-2}\mathbf{A}_k\mathbf{P}_{k+1|N}\mathbf{A}_k^T \end{aligned} \quad (69)$$

and given the definition of $\mathbf{\Omega}$ in (68) results in

$$\begin{aligned}\mathbf{\Omega}=&\int \frac{v_3+\Lambda_k^2}{v_3+n}(\mathbf{P}_{k|k}-\mathbf{A}_k\mathbf{P}_{k+1|k}\mathbf{A}_k^T)\times\\&\mathrm{St}(\mathbf{x}_{k+1};\hat{\mathbf{x}}_{k+1|N},\mathbf{P}_{k+1|N},v_3)d\mathbf{x}_{k+1}\\=&\frac{v_3+\eta_k}{v_3+n}(\mathbf{P}_{k|k}-\mathbf{A}_k\mathbf{P}_{k+1|k}\mathbf{A}_k^T),\end{aligned}\quad(70)$$

where

$$\begin{aligned}\eta_k=&\int \Lambda_k^2\mathrm{St}(\mathbf{x}_{k+1};\hat{\mathbf{x}}_{k+1|N},\mathbf{P}_{k+1|N},v_3)d\mathbf{x}_{k+1}\\=&\int \mathrm{tr}[\mathbf{P}_{k+1|k}^{-1}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1|k})^T]\times\\&\mathrm{St}(\mathbf{x}_{k+1};\hat{\mathbf{x}}_{k+1|N},\mathbf{P}_{k+1|N},v_3)d\mathbf{x}_{k+1}\\=&\mathrm{tr}[\mathbf{P}_{k+1|k}^{-1}\int(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1|k})^T\times\\&\mathrm{St}(\mathbf{x}_{k+1};\hat{\mathbf{x}}_{k+1|N},\mathbf{P}_{k+1|N},v_3)d\mathbf{x}_{k+1}]\\=&\mathrm{tr}\{\mathbf{P}_{k+1|k}^{-1}[\frac{v_3}{v_3-2}\mathbf{P}_{k+1|N}+(\hat{\mathbf{x}}_{k+1|N}-\hat{\mathbf{x}}_{k+1|k})\times\\&(\hat{\mathbf{x}}_{k+1|N}-\hat{\mathbf{x}}_{k+1|k})^T]\}.\end{aligned}\quad(71)$$

Substituting (68)-(71) in (67), the scale matrix $\mathbf{P}_{k|N}$ of the smoothing PDF $p(\mathbf{x}_k|\mathbf{Z}_N)$ can be computed as

$$\begin{aligned}\mathbf{P}_{k|N}=&\mathbf{A}_k\mathbf{P}_{k+1|N}\mathbf{A}_k^T+\frac{(v_3-2)v_3^{''}(v_3+\eta_k)}{v_3(v_3^{''}-2)(v_3+n)}\times\\&(\mathbf{P}_{k|k}-\mathbf{A}_k\mathbf{P}_{k+1|k}\mathbf{A}_k^T).\end{aligned}\quad(72)$$

The proposed Student's t based approximate smoother operates by combining the analytical computations in (32)-(38), (60), (63), (66) and (71)-(72) with the Student's t weighted integrals in (23)-(24), (28)-(30) and (58).

The proposed smoother converges to an existing GA smoother [21] when $v_1=v_2=v_3=v\to+\infty$, which is shown as follows. If $v_1=v_2=v_3=v\to+\infty$, we have

$$\frac{v_3-2}{v_3}\to1\qquad\frac{(v_3-2)v_3^{''}(v_3+\eta_k)}{v_3(v_3^{''}-2)(v_3+n)}\to1.\quad(73)$$

Employing (54) and (73) in (58), we obtain

$$\begin{aligned}\lim_{v\to+\infty}\mathbf{P}_{k,k+1|k}=&\int\mathbf{x}_k\mathbf{f}_k^T(\mathbf{x}_k)N(\mathbf{x}_k;\hat{\mathbf{x}}_{k|k}^{GA},\mathbf{P}_{k|k}^{GA})d\mathbf{x}_k-\\&\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k+1|k}^T.\end{aligned}\quad(74)$$

Using (73) in (72) results in

$$\lim_{v\to+\infty}\mathbf{P}_{k|N}=\mathbf{P}_{k|k}-\mathbf{A}_k(\mathbf{P}_{k+1|k}-\mathbf{P}_{k+1|N})\mathbf{A}_k^T.\quad(75)$$

It is seen from (63), (66) and (74)-(75) that the smoothing estimate vector $\hat{\mathbf{x}}_{k|N}$ and scale matrix $\mathbf{P}_{k|N}$ of the proposed Student's t based approximate smoother will converge to the smoothing estimate vector $\hat{\mathbf{x}}_{k|N}^{GA}$ and the corresponding estimate error covariance matrix $\mathbf{P}_{k|N}^{GA}$ of the existing GA smoother as the dof parameters $v_1=v_2=v_3=v\to+\infty$. Considering that the Student's t PDF converges to the Gaussian PDF as the dof parameters tend to infinity, we obtain

$$\begin{aligned}\lim_{v\to+\infty}p(\mathbf{x}_k|\mathbf{Z}_N)=&\lim_{v\to+\infty}\mathrm{St}(\mathbf{x}_k;\hat{\mathbf{x}}_{k|N},\mathbf{P}_{k|N},v)\\=&N(\mathbf{x}_k;\hat{\mathbf{x}}_{k|N}^{GA},\mathbf{P}_{k|N}^{GA}).\end{aligned}\quad(76)$$

Thus, the existing GA smoother can be deemed as a special case of the proposed Student's t based approximate smoother when $v_1=v_2=v_3=v\to+\infty$.

Remark 1: For general nonlinear systems, closed form solution of the posterior PDF is not available, and approximate approaches have to be employed to calculate the posterior PDF, such as existing Gaussian approximation and the proposed Student's t approximation. These approximations to the posterior PDF all impose a bias on the state estimation. In the proposed Student's t filter, the moment matching approach is used to prevent the growth of the dof parameter, which may induce a further bias on the state estimation. However, if other features of the state estimate are more important to the user than unbiasedness (e.g. root-mean square error (RMSE) and robustness), then the proposed Student's t filter and smoother are attractive. It is shown later in the simulation that the proposed Student's t filter and smoother outperforms existing filters and smoothers in terms of RMSE and robustness, and the proposed Student's t filter and smoother exhibit satisfactory bias.

V. The Computation of the Student's t Weighted Integral

To implement the proposed Student's t based approximate filter and smoother, we need to compute the Student's t weighted integrals whose integrands are all of the following form

$$\mathbf{I}(\mathbf{g})=\int\mathbf{g}(\mathbf{x})\mathrm{St}(\mathbf{x};\mu,\mathbf{\Sigma},v)d\mathbf{x},\quad(77)$$

where $\mathbf{g}(\mathbf{x})$ is a continuous nonlinear function such that the integral in (77) exists.

In this work, the Student's t weighted integral in (77) is approximately computed by the UT since the UT can be used to approximate arbitrary PDF and the first two moment information can be retained. Considering that the Student's t PDF $\mathrm{St}(\mathbf{x};\mu,\mathbf{\Sigma},v)$ has mean vector μ and covariance matrix $\frac{v}{v-2}\mathbf{\Sigma}$, thus by using the UT the Student's t PDF $\mathrm{St}(\mathbf{x};\mu,\mathbf{\Sigma},v)$ can be approximated as follows [8]

$$\mathrm{St}(\mathbf{x};\mu,\mathbf{\Sigma},v)=\sum_{i=0}^{2n}\omega_i\delta(\mathbf{x}-\mathbf{x}_i),\quad(78)$$

where $\delta(\cdot)$ is the Kronecker delta function, $\mathbf{x}_i$ and ω_i are deterministic sigma points and corresponding weights which are given by

$$\begin{cases}\mathbf{x}_i=\mu & \omega_i=\kappa/(n+\kappa)\quad i=0\\\mathbf{x}_i=\mu+\sqrt{\frac{v(n+\kappa)}{v-2}\mathbf{\Sigma}}\mathbf{e}_i & \omega_i=0.5/(n+\kappa)\ \ i=1,\cdots,n\\\mathbf{x}_i=\mu-\sqrt{\frac{v(n+\kappa)}{v-2}\mathbf{\Sigma}}\mathbf{e}_i & \omega_i=0.5/(n+\kappa)\ \ i=n+1,\cdots,2n\end{cases},\quad(79)$$

where κ is the free parameter of the UT, $\sqrt{\mathbf{\Sigma}}$ is the square-root matrix of $\mathbf{\Sigma}$, i.e., $\mathbf{\Sigma}=\sqrt{\mathbf{\Sigma}}\sqrt{\mathbf{\Sigma}}^T$, and $\mathbf{e}_i$ denotes the i-th column vector of a unit matrix.

Substituting (78) in (77), the Student's t weighted integral can be approximated as

$$\mathbf{I}(\mathbf{g})=\int\mathbf{g}(\mathbf{x})\left[\sum_{i=0}^{2n}\omega_i\delta(\mathbf{x}-\mathbf{x}_i)\right]d\mathbf{x}=\sum_{i=0}^{2n}\omega_i\mathbf{g}(\mathbf{x}_i),\quad(80)$$

where $\mathbf{x}_i$ and ω_i are given in (79).

By using the UT in (79)-(80) to compute the Student's t weighted integrals in (23)-(24), (28)-(30) and (58), a new unscented Student's t based approximate filter and unscented Student's t based approximate smoother can be developed respectively.

VI. Simulation

In this simulation, the superior performance of the proposed methods as compared with existing methods is shown in the problem of tracking an agile target in two dimensional space executing a manoeuvring turn with unknown and time-varying turn rate. This problem was used to illustrate the performance of the CKF [5] and GA smoother [21]. The process and measurement models are formulated as follows

$$\mathbf{x}_k = \begin{bmatrix} 1 & \frac{\sin\Omega T_0}{\Omega} & 0 & \frac{\cos\Omega T_0 - 1}{\Omega} & 0 \\ 0 & \cos\Omega T_0 & 0 & -\sin\Omega T_0 & 0 \\ 0 & \frac{1-\cos\Omega T_0}{\Omega} & 1 & \frac{\sin\Omega T_0}{\Omega} & 0 \\ 0 & \sin\Omega T_0 & 0 & \cos\Omega T_0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{81}$$

$$\mathbf{z}_k = \begin{bmatrix} r_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} \sqrt{\varsigma_k^2 + \eta_k^2} \\ \tan^{-1}(\frac{\eta_k}{\varsigma_k}) \end{bmatrix} + \mathbf{v}_k, \tag{82}$$

where the state vector $\mathbf{x} = [\varsigma \ \dot{\varsigma} \ \eta \ \dot{\eta} \ \Omega]^T$; ς and η denote positions, $\dot{\varsigma}$ and $\dot{\eta}$ denote velocities in the x and y directions respectively, Ω denotes constant but unknown turn rate, and T_0 denotes the time-interval between two consecutive measurements. Similar to [14], outlier corrupted process and measurement noises are generated according to

$$\mathbf{w}_k \sim \begin{cases} N(\mathbf{0}, \mathbf{Q}) & \text{w.p. } 0.95 \\ N(\mathbf{0}, 50\mathbf{Q}) & \text{w.p. } 0.05 \end{cases} \tag{83}$$

$$\mathbf{v}_k \sim \begin{cases} N(\mathbf{0}, \mathbf{R}) & \text{w.p. } 0.95 \\ N(\mathbf{0}, 100\mathbf{R}) & \text{w.p. } 0.05 \end{cases}, \tag{84}$$

where w.p. denotes "with probability". Equations (83)-(84) mean that $\mathbf{w}_k$ and $\mathbf{v}_k$ are most of the times drawn from a Gaussian distribution with nominal covariance matrix $\mathbf{Q}$ or $\mathbf{R}$ and five percent of process and measurement noise values are generated from Gaussian distributions with severely increased covariance matrix. Process and measurement noises, which are generated in terms of (83)-(84), have heavy tails. In this simulation, parameters T_0, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{x}_0$ and $\mathbf{P}_{0|0}$ are the same as defined in [5].

In this simulation, the existing UKF [8], the outlier robust UKF [18], the Student's t extended Kalman filter (EKF) [14], the UKS [9], [21], the outlier robust UKS [18], and the proposed unscented Student's t based approximate filter and smoother are tested. In the proposed and existing methods, the free parameter is chosen as $\kappa = 0$ to ensure the numerical stability of the UT [5] and the dof parameters are set $v_1 = v_2 = v_3 = 3$ as suggested in [14]. To compare the performance of existing methods and the proposed methods, the RMSEs of position, velocity and turn rate, which were defined in [5], are chosen as performance metrics.

Fig. 1 shows the true and estimated trajectories of an aircraft obtained from existing methods and the proposed methods in a single Monte Carlo run. Fig. 2–Fig. 4 show the RMSEs of position, velocity and turn rate, respectively from the existing methods and the proposed methods by making 250 independent Monte Carlo runs. Table I shows the averaged RMSEs of the proposed methods and existing methods over the last 20s. It is seen from Fig. 1 that the estimated trajectories from the proposed methods are closer to the true trajectory as compared with existing methods. We can see from Fig. 2–Fig. 4 and Table I that the proposed filter has considerably improved filtering accuracy as compared to the existing UKF, the outlier robust UKF and the Student's t EKF. And the proposed smoother has increased smoothing accuracy as compared to the existing UKS and outlier robust UKS.

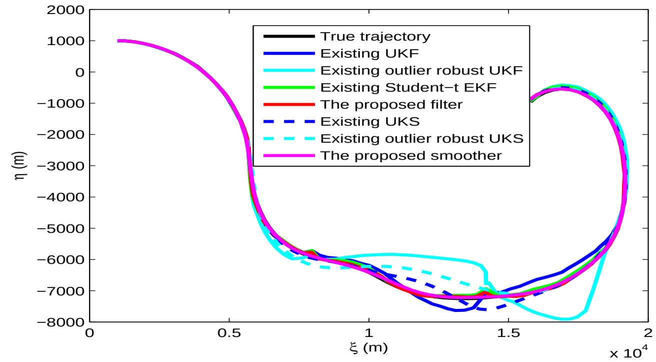

Fig. 1: True and estimated trajectories of the target

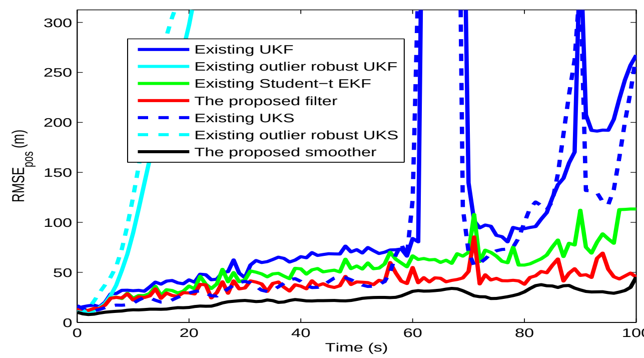

Fig. 2: RMSEs of the position of the target

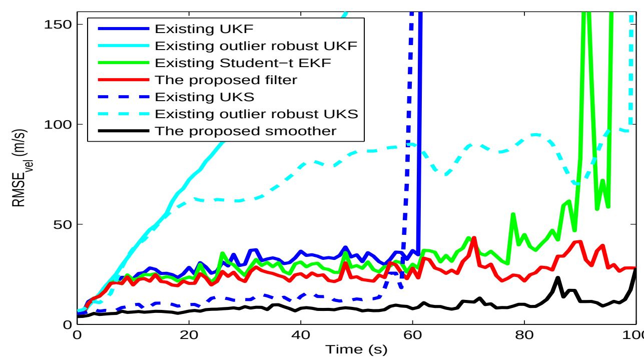

Fig. 3: RMSEs of the velocity of the target

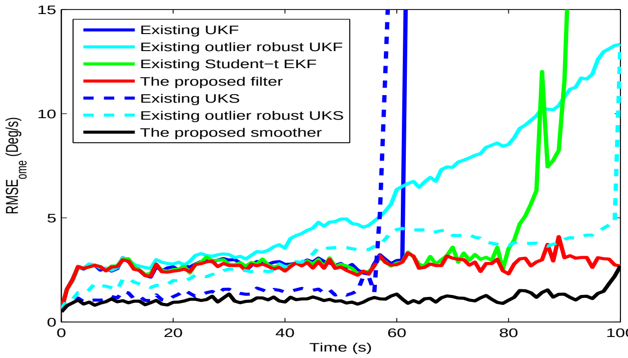

Fig. 4: RMSEs of the turn rate of the target

TABLE I: Averaged RMSEs of the proposed methods and existing methods over the last 20s

Estimator	position (m)	velocity (m/s)	turn rate
Existing UKF [8]	184.5	10573.9	48.8
Existing outlier robust UKF [18]	194.5	351.7	11.1
Existing Student's EKF [14]	83.6	68.9	17.3
The proposed filter	50.6	31.4	3.0
Existing UKS [9]	169.1	10783.2	49.5
Existing outlier robust UKS [18]	491.3	108.5	4.5
The proposed smoother	34.0	13.6	1.4

Moreover, as expected, the proposed smoother has higher estimation accuracy than the proposed filter.

The RMSEs from the existing UKF and UKS increase abruptly after 60s since they are specially designed for Gaussian process and measurement noises so that they are sensitive to process and measurement outliers. The RMSEs of the velocity and the turn rate from the existing Student's t EKF also increase abruptly after 80s, which is incurred by the truncation errors of the first-order linearisation. Moreover, the RMSEs from the existing outlier robust UKF and outlier robust UKS diverge since they assume a well behaved process noise so that they are sensitive to process outliers.

To evaluate the consistencies of the proposed filter and smoother, the filter-estimated RMSE and smoother-estimated RMSE are respectively compared with the true filter RMSE and smoother RMSE, where the filter-estimated RMSE and smoother-estimated RMSE are respectively the square-root of the averaged (over Monte Carlo runs) appropriate diagonal entries of the filtering and smoothing covariance matrices [5]. The state estimate is judged to be consistent if the estimated RMSE equals to the true RMSE [5], [7].

Fig. 5–Fig. 7 show the filter-estimated RMSEs, the smoother-estimated RMSEs, the true filter RMSEs and smoother RMSEs of position, velocity and turn rate from the proposed filter and smoother by making 1000 independent Monte Carlo runs. We can see from Fig. 5–Fig. 7 that there are small differences between the filter-estimated RMSEs and the true filter RMSEs, and between the smoother-estimated RMSEs and the true smoother RMSEs, which are induced by the Student's t approximation to the posterior filtering and smoothing PDFs. However, the filter-estimated RMSEs and smoother-estimated RMSEs closely follow the trends of the

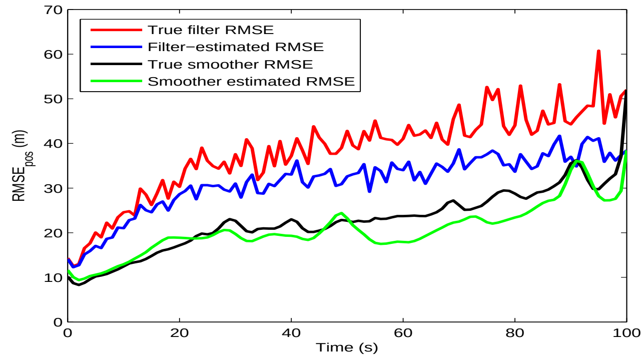

Fig. 5: True and estimated RMSEs of the position from the proposed filter and smoother

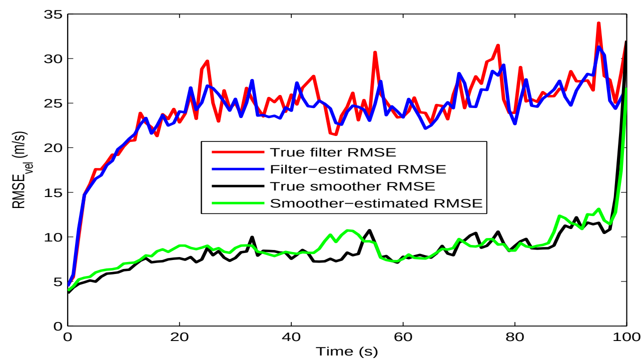

Fig. 6: True and estimated RMSEs of the velocity from the proposed filter and smoother

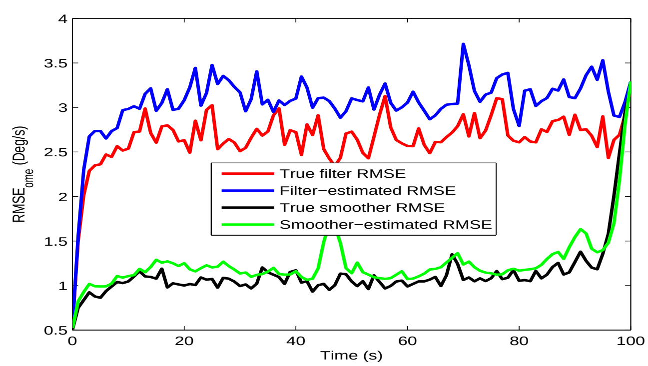

Fig. 7: True and estimated RMSEs of the turn rate from the proposed filter and smoother

true filter RMSEs and the true smoother RMSEs respectively. Thus, the proposed Student's t filter and smoother exhibit acceptable consistencies.

To evaluate the unbiasedness of the proposed filter and smoother, the biases of position, velocity and turn rate from the proposed filter and smoother are calculated by making 1000 independent Monte Carlo runs, which are shown in Fig. 8–Fig. 10. We can see from Fig. 8–Fig. 10 that the biases of position, velocity and turn rate from the proposed filter and

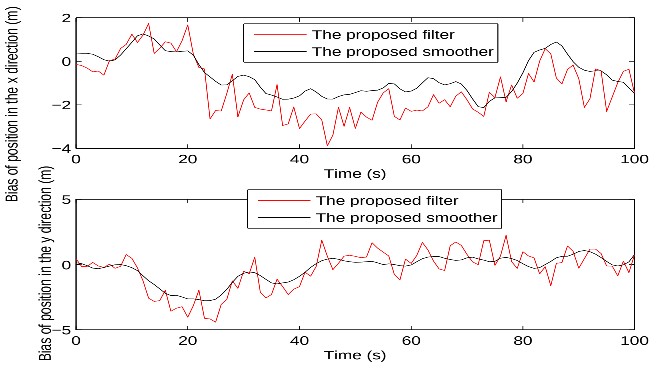

Fig. 8: Biases of the position from the proposed filter and smoother

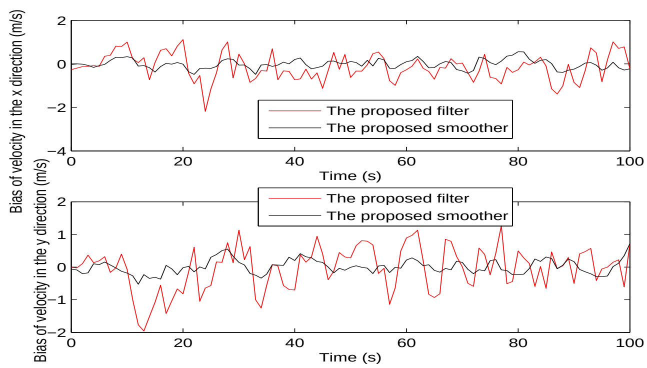

Fig. 9: Biases of the velocity from the proposed filter and smoother

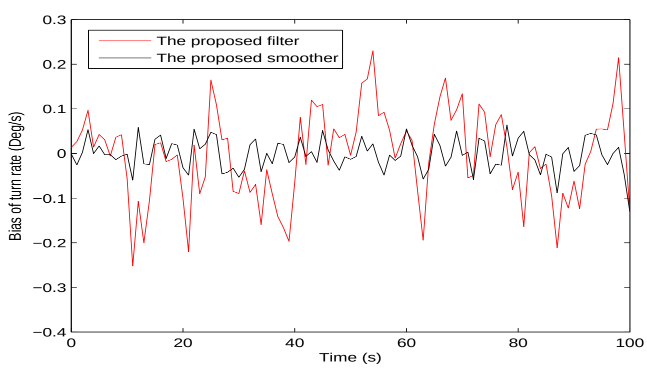

Fig. 10: Biases of the turn rate from the proposed filter and smoother

smoother are all non zero. Consequently, the proposed filter and smoother are biased, but we emphasize that all existing nonlinear filters and smoothers diverge after 80s in Fig. 2–Fig. 4. Thus, the proposed filter and smoother are more robust and exhibit smaller biases as compared with existing filters and smoothers.

VII. Conclusions

In this paper, novel Student's t based approaches to approximate a filter and smoother are proposed, which require analytical computations and Student's t weighted integrals. The moment matching approach is used to constrain the increase of dof parameters of the posterior filtering and smoothing PDFs, and the UT is used to compute the Student's t weighted integrals involved in the proposed Student's t based approximate filter and smoother. Simulation results illustrate the proposed methods considerably outperform existing methods in an application with heavy tailed process and measurement noises.

References

[1] M-J. Yu. "INS/GPS integration system using adaptive filter for estimating measurement noise variance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1786–1792, Apr 2012.

[2] B. Jia and M. Xin. "Sparse-grid quadrature H_∞ filter for discrete-time systems with uncertain noise statistics," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1626–1636, July 2013.

[3] Y. G. Zhang, Y. L. Huang, N. Li and L. Zhao. "Interpolatory cubature Kalman filters," *IET Control Theory & Applications*, vol. 9, no. 11, pp. 1731–1739, Sep 2015.

[4] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 4, pp. 910−927, May 2000.

[5] I. Arasaratnam and S. Haykin, "Cubature Kalman filter," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254−1269, Jun 2009.

[6] X. X. Wang, Y. Liang, Q. Pan and C. H. Zhao, "Gaussian filter for nonlinear systems with one-step randomly delayed measurements," *Automatica*, vol. 49, no. 4, pp. 976−986, Apr 2013.

[7] Y. L. Huang, Y. G. Zhang, X. X. Wang and L. Zhao. "Gaussian filter for nonlinear systems with correlated noises at the same epoch," *Automatica*, vol. 60, no. 10, pp. 122-126, Oct, 2015.

[8] S. J. Julier and J. K. Uhlman, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401−422, Mar 2004.

[9] S. Särkkä, "Unscented Rauch-Tung-Striebel smoother," *IEEE Transactions on Automatic Control*, vol. 53, no. 3, pp. 845−849, Apr 2008.

[10] B. Jia, M. Xin, and Y. Cheng, "High-degree cubature Kalman filter," *Automatica*, vol. 49, no. 2, pp. 510−518, Feb 2013.

[11] I. Arasaratnam and S. Haykin, "Cubature Kalman smoothers," *Automatica*, vol. 47, no. 10, pp. 2245−2250, Oct 2011.

[12] B. Jia and M. Xin, "A new class of nonlinear Rauch-Tung-Striebel cubature Kalman smoothers," *ISA Transactions*, vol. 55, no. 3, pp. 72–80, Mar 2015.

[13] Y. G. Zhang, Y. L. Huang, N. Li and L. Zhao. "Embedded cubature Kalman filter with adaptive setting of free parameter," *Signal Processing*, vol. 114, no. 9, pp. 112–116, Sep 2015.

[14] M. Roth, E. Özkan, and F. Gustafsson, "A Student's filter for heavy tailed process and measurement noise," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 5770–5774.

[15] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "An outlier-robust Kalman filter," *in IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1551–1558, May 2011.

[16] G. Agamennoni, J.I. Nieto, and E.M. Nebot,"Approximate inference in state-space models with heavy-tailed noise," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5024–5037, Oct. 2012.

[17] H. Nurminen, T. Ardeshiri, R. Piché, and F. Gustafsson, "Robust inference for state-space models with skewed measurement noise," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1898–1902, Nov. 2015.

[18] R. Piché, S. Särkkä, and J. Hartikainen, "Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate Student-t distribution," *in Proceedings of MLSP*, Sept. 2012.

[19] A. Jepson, D. Fleet, and T. El-Maraghi, "Robust on-line appearance models for visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intellgence*, vol. 25, no. 10, pp. 1296–1311, Oct. 2003.

[20] J. Loxam and T. Drummond, "Student mixture filter for robust, real-time visual tracking," *in Proceedings of 10th European Conference on Computer Vision: Part III*, 2008.

[21] S. Särkkä, and J. Hartikainen, "On Gaussian optimal smoothing of non-linear state space models," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1938–1941, Aug 2010.