Research Article

A Novel Robust Student's t -Based Cubature Information Filter with Heavy-Tailed Noises

Yongtao Shui,¹ Xiaogang Wang¹,¹ Wutao Qin,² Yu Wang,¹ Baojun Pang,¹ and Naigang Cui¹

¹School of astronautics, Harbin Institute of Technology, Harbin, China

²Nanjing Research Institute of Electronics Technology, Nanjing, China

Correspondence should be addressed to Xiaogang Wang; wangxiaogang@hit.edu.cn

Received 5 January 2020; Revised 25 February 2020; Accepted 2 July 2020; Published 17 July 2020

Academic Editor: Giovanni Palmerini

Copyright © 2020 Yongtao Shui et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a novel robust Student's t -based cubature information filter is proposed for a nonlinear multisensor system with heavy-tailed process and measurement noises. At first, the predictive probability density function (PDF) and the likelihood PDF are approximated as two different Student's t distributions. To avoid the process uncertainty induced by the heavy-tailed process noise, the scale matrix of the predictive PDF is modeled as an inverse Wishart distribution and estimated dynamically. Then, the predictive PDF and the likelihood PDF are transformed into a hierarchical Gaussian form to obtain the approximate solution of posterior PDF. Based on the variational Bayesian approximation method, the posterior PDF is approximated iteratively by minimizing the Kullback-Leibler divergence function. Based on the posterior PDF of the auxiliary parameters, the predicted covariance and measurement noise covariance are modified. And then the information matrix and information state are updated by summing the local information contributions, which are computed based on the modified covariance. Finally, the state, scale matrix, and posterior densities are estimated after fixed point iterations. And the simulation results for a target tracking example demonstrate the superiority of the proposed filter.

1. Introduction

To obtain the reliable and precise information, multisensor systems have become more and more popular in a wide range of applications such as cooperative tracking, autonomous navigation, signal processing, and guidance [1, 2]. Therefore, how to best fuse the observations from different sensors has gained much attention in several decades. Kalman filter (KF), one of the most popular state estimation methods, has been used extensively in many multisensor estimation problems [3, 4]. However, the KF is a linear state estimation method, and it cannot get an optimal estimation result with non-Gaussian process and measurement noises. But in many engineering applications, such as tracking agile target in clutter with bearing sensors, the system is nonlinear and the heavy-tailed noises are presented [5, 6]. In this case, the performance of the classical KF degrades dramatically.

To apply the KF in nonlinear system and improve the accuracy of nonlinear approximation, the first order linearization-based extended Kalman filter (EKF) [7], the

unscented transformation-based unscented Kalman filtering (UKF) [8, 9], quadrature rule-based Gauss-Hermite filtering (GHF) [10], polynomial interpolation-based divided difference filtering (DDF) [11], and the spherical-radical rule-based cubature Kalman filter (CKF) [12, 13] have been proposed successively. Although the information filter (IF) is an algebraically equivalent form of Kalman filter, it is more suitable for multisensor estimation, since the IF can reduce the computational burden significantly and be decentralized easily [14]. Soon, the abovementioned filters were formulated in the information type and applied to the multisensor estimation, such as extended information filter (EIF) [15], unscented information filter (UIF) [16], and cubature information filter (CIF) [17]. Unfortunately, the performance of these information filters will decrease in the presence of heavy-tailed noises, since they all derived based on the Gaussian assumption.

To tackle the heavy-tailed non-Gaussian noise, some Huber-based information filters have been proposed by minimizing a cost function which is a combined l_1 and l_2 norm

[18], such as Huber-based UIF [19] and Huber-based CIF [20]. Nevertheless, the Huber's weight function cannot converge to zero even as the error approaches infinity. It indicates the measurements with large errors still will be employed, and the estimation performance will be deteriorated [5, 21]. Besides Huber's technique, some maximum correntropy (MC) information filters such as MC-UIF [22] and MC-CIF [5] have been proposed by employing the maximum correntropy criterion as the optimality criterion. However, the estimation covariance cannot be developed due to the lack of theoretical basis, which will degrade the performance of MC Kalman filters.

Since the Student's t distribution is a reasonable model of the heavy-tailed noise, some Student's t -based Kalman filters (STKF) have been proposed by modeling the heavy-tailed noises as the Student's t distribution [23–27]. To obtain the approximate solution of posterior estimations, the posterior probability density function (PDF) has also been approximated by the Student's t distribution [21, 23], in which the growth of the dof parameter in STKF degrades the estimation accuracy obviously. To tackle with the difficulty in choosing scale matrix and dof parameter, variational Bayesian (VB) approach is employed to jointly estimate the state vector, scale matrix, and dof parameter [24, 25]. In [28], the inaccurate process noise covariance noise and measurement noise covariance matrix is estimated by choosing inverse Wishart priors. In [29], both skewed noise and heavy-tailed noise are considered, whereas Gaussian scale mixture distributions are used to model one-step prediction and likelihood PDFs. Despite the excellent work mentioned above, there have been relatively few results concerning the nonlinear multisensor estimation problem with heavy-tailed process, and measurement noises have not been considered, which serves as the main motivation of this paper.

To cope with the nonlinear multisensor estimation problem with heavy-tailed process and measurement noises, a novel robust Student's t -based cubature information filter (NRSTCIF) is proposed in this paper. At first, the one-step predicted state vector and nominal covariance is propagated based on the cubature rule. Then, the predictive PDF and the likelihood PDF are transformed into a hierarchical Gaussian form to obtain the approximate solution of posterior PDF. Next, the posterior densities are approximated iteratively based on the variational Bayesian (VB) approximation method, where the information matrix and information state are updated by summing the local information contributions. Finally, the state, scale matrix, and posterior densities are estimated after fixed point iterations.

The rest of the paper is organized as follows. In Sec. II, the problem is formulated. Sec. III gives the derivation of the proposed filter. The simulations are conducted in Sec. IV, and the conclusions are drawn in Sec. V.

2. Problem Formulation

Consider a discrete-time stochastic nonlinear system described by

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

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \quad (2)$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the n -dimensional state, $\mathbf{z}_k \in \mathbb{R}^m$ is the m -dimensional measurement, and $\mathbf{w}_k \in \mathbb{R}^n$ and $\mathbf{v}_k \in \mathbb{R}^m$ are zero mean heavily tailed process noise and measurement noise with nominal covariance $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively.

Since $\mathbf{v}_k$ is a heavy-tailed measurement noise, its distribution is approximated as

$$p(\mathbf{v}_k) = \text{St}(\mathbf{v}_k; \mathbf{0}, \mathbf{R}_k, \nu) \quad (3)$$

where $\text{St}(\mathbf{v}_k; \mathbf{0}, \mathbf{R}_k, \nu)$ denotes the Student's t probability distribution function (PDF) of $\mathbf{v}_k$, with mean $\mathbf{0}$ and scale matrix $\mathbf{R}_k$, ν is the dof parameter.

Based on the measurement equation and the distribution of measurement noise, the likelihood PDF $p(\mathbf{z}_k|\mathbf{x}_k)$ can be formulated as

$$p(\mathbf{z}_k|\mathbf{x}_k) = \text{St}(\mathbf{z}_k; \mathbf{h}(\mathbf{x}_k), \mathbf{R}_k, \nu) \quad (4)$$

where $\mathbf{h}(\mathbf{x}_k)$ is the mean vector of $\mathbf{z}_k$.

Due to the heavy-tailed process noise, the predictive PDF $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ is also approximated as a Student's t distribution.

$$p(\mathbf{x}_k|\mathbf{Z}_{k-1}) = \text{St}(\mathbf{x}_k; \mathbf{u}\hat{\mathbf{x}}_{k|k-1}, \mathcal{T}_k, \mu) \quad (5)$$

where $\hat{\mathbf{x}}_{k|k-1}$ is the mean vector, μ is the dof parameter, and $\mathcal{T}_k$ is the scale matrix which will be approximated as the inverse Wishart distribution in the next section.

3. Derivation of the Proportional Filter

3.1. Brief Review of Extended Information Filter. Based on the state model given by (1), the predicted information matrix $\mathbf{Y}_{k|k-1}$ and information state $\hat{\mathbf{y}}_{k|k-1}$ can be formulated as [16, 17].

$$\begin{aligned} \mathbf{Y}_{k|k-1} &= \mathbf{P}_{k|k-1}^{-1} = [\Phi \mathbf{Y}_{k-1|k-1}^{-1} \Phi^T + \mathbf{Q}_{k-1}]^{-1} \\ \hat{\mathbf{y}}_{k|k-1} &= \mathbf{Y}_{k|k-1} \hat{\mathbf{x}}_{k|k-1} = \mathbf{Y}_{k-1|k-1} \mathbf{f}(\hat{\mathbf{x}}_{k-1|k-1}) \end{aligned} \quad (6)$$

where Φ is the state-transition matrix which can be computed based on (1).

Then, the updated information matrix $\mathbf{Y}_{k|k}$ and information state $\hat{\mathbf{y}}_{k|k}$ can be formulated as

$$\begin{aligned} \hat{\mathbf{y}}_{k|k} &= \hat{\mathbf{y}}_{k|k-1} + \mathbf{i}_k \\ \mathbf{Y}_{k|k} &= \mathbf{Y}_{k|k-1} + \mathbf{I}_k \end{aligned} \quad (7)$$

where $\mathbf{i}_k$ and $\mathbf{I}_k$ are information state contribution and information matrix contribution, respectively. They are given by

$$\mathbf{i}_k = \mathbf{H}_k^T \mathbf{R}_k^{-1} [\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}] \quad (8)$$

$$\mathbf{I}_k = \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \quad (9)$$

where $\mathbf{z}_k$ is the measurement at time k and $\mathbf{H}_k$ is the measurement matrix.

3.2. *The Transformation of the Predictive PDF and the Likelihood PDF.* To get the approximate solution of posterior PDF $p(\mathbf{x}_k|\mathbf{Z}_k)$, the predictive PDF $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ and the likelihood PDF $p(\mathbf{z}_k|\mathbf{x}_k)$ given by (5) and (4) are written as an infinite mixture of Gaussian distribution [24], i.e.,

$$p(\mathbf{x}_k|\mathbf{Z}_{k-1}) = \int N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathcal{T}_k/\varsigma_k) \mathcal{G}(\varsigma_k; \frac{\mu}{2}, \frac{\mu}{2}) d\varsigma_k \quad (10)$$

$$p(\mathbf{z}_k|\mathbf{x}_k) = \int N(\mathbf{x}_k; \mathbf{h}(\mathbf{x}_k), \mathbf{R}_k/\eta_k) \mathcal{G}(\eta_k; \frac{v}{2}, \frac{v}{2}) d\eta_k \quad (11)$$

where ς_k and η_k are auxiliary parameters, $\mathcal{G}(\omega; a, b)$ is the Gamma PDF of auxiliary parameter ω . $\mathcal{G}(\omega; a, b)$ is defined by

$$\mathcal{G}(\omega; a, b) = \frac{b^a}{\Gamma(a)} \omega^{a-1} e^{-b\omega} \quad (12)$$

where $\Gamma(\cdot)$ denotes the Gamma function. Then, $\ln \mathcal{G}(\omega; a, b)$ can be formulated as

$$\begin{aligned} \ln \mathcal{G}(\omega; a, b) &= \ln \omega^{a-1} + \ln e^{-b\omega} + \ln \frac{b^a}{\Gamma(a)} \\ &= (a-1) \ln \omega - b\omega + c_{a,b} \end{aligned} \quad (13)$$

Since the process noise is a zero mean vector with nominal covariance $\mathbf{Q}_k$, the state vector and its associated nominal covariance can be propagated by

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \frac{1}{2n} \cdot \sum_{i=1}^{2n} \mathbf{f}(\chi_{i,k-1}) \\ \mathbf{P}_{k|k-1} &= \frac{1}{2n} \cdot \sum_{i=1}^{2n} (\mathbf{f}(\chi_{i,k-1}) - \hat{\mathbf{x}}_{k|k-1})(\mathbf{f}(\chi_{i,k-1}) \\ &\quad - \hat{\mathbf{x}}_{k|k-1})^T + \mathbf{Q}_{k-1} \end{aligned} \quad (14)$$

where $\chi_{i,k-1}$ ($i = 1, 2, \dots, 2n$) is the i -th cubature point generated by

$$\chi_{i,k-1} = \mathbf{x}_{k-1} + \mathbf{S}_{k-1} \boldsymbol{\xi}_i, \quad i = 1, 2, \dots, 2n \quad (15)$$

where $\mathbf{S}_{k-1}$ is the Cholesky decomposition of $\mathbf{P}_{k-1}$ satisfying $\mathbf{P}_{k-1} = \mathbf{S}_{k-1} \cdot \mathbf{S}_{k-1}^T$, and $\boldsymbol{\xi}_i$ is the i -th element of the following set

$$\sqrt{n} \left[\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -1 \end{pmatrix} \right] \quad (16)$$

To accurately describe the predictive PDF, the VB approximation method is applied to dynamically estimate the scale matrix $\mathcal{T}_k$ rather than set nominal covariance $\mathbf{P}_{k|k-1}$ as $\mathcal{T}_k$. Then, the PDF of $\mathcal{T}_k$ can be appropriated as

$$\begin{aligned} p(\mathcal{T}_k) &= \text{IW}(\mathcal{T}_k; t_k, \mathbf{T}_k) \\ &= \frac{|\mathbf{T}_k|^{t_k/2} |\mathcal{T}_k|^{-(t_k+n+1)/2} \exp[-0.5 \text{tr}(\mathbf{T}_k \mathcal{T}_k^{-1})]}{2^{n t_k/2} \Gamma_n(t_k/2)} \end{aligned} \quad (17)$$

where $\text{IW}(\mathcal{T}_k; t_k, \mathbf{T}_k)$ denotes the inverse Wishart distribution, whose functional form will not change between prior and posterior, t_k is the dof parameter, and $\mathbf{T}_k$ is inverse scale matrix. $\Gamma_n(\cdot)$ and $\text{tr}(\cdot)$ denote the n -variate gamma function and the trace operation, respectively.

Then, $\ln [\text{IW}(\mathcal{T}_k; t_k, \mathbf{T}_k)]$ can be formulated as

$$\begin{aligned} \ln [\text{IW}(\mathcal{T}_k; t_k, \mathbf{T}_k)] &= \ln \left(|\mathcal{T}_k|^{-(t_k+n+1)/2} \right) \\ &\quad + \ln \left(\exp[-0.5 \text{tr}(\mathbf{T}_k \mathcal{T}_k^{-1})] \right) \\ &\quad + \left\{ \ln \left(|\mathbf{T}_k|^{t_k/2} \right) - \ln \left[2^{n t_k/2} \Gamma_n(t_k/2) \right] \right\} \\ &= -\frac{t_k + n + 1}{2} \ln |\mathcal{T}_k| - \frac{\text{tr}(\mathbf{T}_k \mathcal{T}_k^{-1})}{2} + c_{\mathcal{T}} \end{aligned} \quad (18)$$

According to the property of inverse Wishart distribution, we have

$$\mathbb{E}[\mathcal{T}_k^{-1}] = (t_k - n - 1) \mathbf{T}_k^{-1} \text{ for } t_k > n + 1 \quad (19)$$

To capture the statistics of $\mathcal{T}_k$, the nominal covariance $\mathbf{P}_{k|k-1}$ is chosen as the mean value of $p(\mathcal{T}_k)$. Exploiting (19), the inverse scale matrix $\mathbf{T}_k$ is given by

$$\mathbf{T}_k = \tau \mathbf{P}_{k|k-1} \quad (20)$$

where the nonnegative tuning parameter $\tau = t_k - n - 1$.

Based on the (10), (11), and (17), the predictive PDF $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ and the likelihood PDF $p(\mathbf{z}_k|\mathbf{x}_k)$ can be transformed as follows:

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{Z}_{k-1}) &= N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathcal{T}_k/\varsigma_k) \\ p(\mathcal{T}_k) &= \text{IW}(\mathcal{T}_k; t_k, \mathbf{T}_k) \end{aligned} \quad (21)$$

$$p(\varsigma_k) = \mathcal{G}\left(\varsigma_k; \frac{\mu}{2}, \frac{\mu}{2}\right)$$

$$\begin{aligned} p(\mathbf{z}_k|\mathbf{x}_k, \eta_k) &= N(\mathbf{x}_k; \mathbf{h}(\mathbf{x}_k), \mathbf{R}_k/\eta_k) \\ p(\eta_k) &= \mathcal{G}\left(\eta_k; \frac{v}{2}, \frac{v}{2}\right) \end{aligned} \quad (22)$$

Note that we consider the case that the heavy-tailed process noise statistics, i.e., the scale matrix $\mathcal{T}_k$, is unknown and time varying, which is usually the case in practical applications especially for target tracking [29]. Therefore, the scale matrix $\mathcal{T}_k$ of process noise is modeled as inverse Wishart distribution and estimated dynamically. For measurement noise, we assume that the noise statistics $\mathbf{R}_k$ is known based on the prior information of sensors; therefore, it can be chosen wisely rather than estimated.

3.3. *Derivation of Measurement Update Based on VB Approximation Method.* To estimate the state $\mathbf{x}_k$, the joint

posterior PDF $p(\mathbf{x}_k, \varsigma_k, \mathcal{T}_k, \eta_k | \mathbf{Z}_k)$ is required to be approximated. By making use of the VB approximation method, we have

$$p(\mathbf{x}_k, \varsigma_k, \mathcal{T}_k, \eta_k | \mathbf{Z}_k) \approx q(\mathbf{x}_k)q(\varsigma_k)q(\mathcal{T}_k)q(\eta_k) \quad (23)$$

where $q(\cdot)$ is the approximate posterior PDF, which could be obtained by minimizing the following function.

$$J = \text{KL}(q(\mathbf{x}_k)q(\varsigma_k)q(\mathcal{T}_k^{-1})q(\eta_k) || p(\mathbf{x}_k, \varsigma_k, \mathcal{T}_k^{-1}, \eta_k | \mathbf{Z}_k)) \quad (24)$$

where KL denotes the Kullback-Leibler divergence function.

Minimizing (24) can reach the following VB-marginal [30].

$$\ln q(\varphi) = \mathbb{E}_{\Theta(-\varphi)} [\ln p(\Theta, \mathbf{Z}_k)] + c_\varphi \quad (25)$$

where $\Theta = [\mathbf{x}_k, \varsigma_k, \mathcal{T}_k, \eta_k]$ is a set consists of $\mathbf{x}_k, \varsigma_k, \mathcal{T}_k$ and η_k , φ is an arbitrary element of Θ and $\Theta(-\varphi)$ denotes the set Θ without φ , and c_φ is the constant related to φ .

Since the PDF given by (21) and (22) are conditional independent, $p(\Theta, \mathbf{Z}_k)$ could be factored as

$$p(\Theta, \mathbf{Z}_k) = p(\mathbf{z}_k | \mathbf{x}_k, \eta_k) p(\mathbf{x}_k | \varsigma_k, \mathbf{Z}_{k-1}) p(\varsigma_k) p(\mathcal{T}_k) p(\eta_k) \quad (26)$$

Substituting (21) and (22) into (26), we have

$$p(\Theta, \mathbf{Z}_k) = N(\mathbf{x}_k; \mathbf{h}(\mathbf{x}_k), \mathbf{R}_k/\eta_k) N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathcal{T}_k/\varsigma_k) \cdot \mathcal{G}\left(\varsigma_k; \frac{\mu}{2}, \frac{\mu}{2}\right) \text{IW}(\mathcal{T}_k; t_k, \mathbf{T}_k) \mathcal{G}\left(\eta_k; \frac{v}{2}, \frac{v}{2}\right) \quad (27)$$

Then, $\ln p(\Theta, \mathbf{Z}_k)$ can be expressed as [24].

$$\begin{aligned} \ln p(\Theta, \mathbf{Z}_k) = & \left(\frac{m+v}{2} - 1\right) \ln \eta_k - \frac{v}{2} \eta_k \\ & - \frac{\eta_k}{2} (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k))^T \mathbf{R}_k^{-1} (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k)) \\ & + \left(\frac{n+\mu}{2} - 1\right) \ln \varsigma_k - \frac{\mu}{2} \varsigma_k \\ & - \frac{\varsigma_k}{2} (\mathbf{x}_k - \mathbf{x}_{k|k-1})^T \mathcal{T}_k^{-1} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \\ & - \frac{\text{tr}(\mathbf{T}_k \mathcal{T}_k^{-1})}{2} - \frac{n+t_k+2}{2} \ln |\mathcal{T}_k| \end{aligned} \quad (28)$$

By updating one element of Θ and keeping the other fixed at their last estimated values, the VB-marginals given by (25) can be solved iteratively.

Let $\phi = \varsigma_k$ and substituting (28) in (25) results in

$$\begin{aligned} \ln q^{(i+1)}(\varsigma_k) = & \left(\frac{n+\mu}{2} - 1\right) \ln \varsigma_k \\ & - \frac{\left(\mu + \text{tr}\left(\mathcal{P}_k^{(i)} \mathbb{E}^{(i)}[\mathcal{T}_k^{-1}]\right)\right)}{2} \varsigma_k + c_\varsigma \end{aligned} \quad (29)$$

where the superscript “(i)” represents the i -th iteration. $\mathcal{P}_k^{(i)}$ is given by

$$\begin{aligned} \mathcal{P}_k^{(i)} = & \mathbb{E}^{(i)} \left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) (\mathbf{x}_k - \mathbf{x}_{k|k-1})^T \right] \\ = & \mathbb{E}^{(i)} \left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}^{(i)}) (\mathbf{x}_k - \mathbf{x}_{k|k}^{(i)})^T \right] \\ & + (\hat{\mathbf{x}}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k-1}) (\mathbf{x}_{k|k}^{(i)} - \mathbf{x}_{k|k-1})^T \\ = & \mathbf{P}_{k|k}^{(i)} + (\hat{\mathbf{x}}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k-1}) (\mathbf{x}_{k|k}^{(i)} - \mathbf{x}_{k|k-1})^T \end{aligned} \quad (30)$$

And the initial value of iteration is set as $\hat{\mathbf{x}}_{k|k}^{(0)} = \hat{\mathbf{x}}_{k|k-1}$, $\mathbf{P}_{k|k}^{(i)} = \mathbf{P}_{k|k-1}$, and $\mathbb{E}^{(0)}[\mathcal{T}_k^{-1}] = (t_k - n - 1) \mathbf{T}_k^{-1}$.

Contrasting (13) with (29), the approximate posterior PDF $q^{(i+1)}(\varsigma_k)$ can be given by

$$q^{(i+1)}(\varsigma_k) = \mathcal{G}\left(\varsigma_k; \alpha_k^{(i+1)}, \beta_k^{(i+1)}\right) \quad (31)$$

where $\alpha_k^{(i+1)} = n + \mu/2$, $\beta_k^{(i+1)} = \mu + \text{tr}(\mathcal{P}_k^{(i)} \mathbb{E}^{(i)}[\mathcal{T}_k^{-1}])/2$.

Let $\phi = \eta_k$ and substituting (28) in (25) results in

$$\ln q^{(i+1)}(\eta_k) = \left(\frac{m+v}{2} - 1\right) \ln \eta_k - \frac{\left(v + \text{tr}(\mathcal{R}_k^{(i)} \mathbf{R}_k^{-1})\right)}{2} \eta_k + c_\eta \quad (32)$$

where $\mathcal{R}_k^{(i)}$ is given by

$$\begin{aligned} \mathcal{R}_k^{(i)} = & \mathbb{E}^{(i)} \left[(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k)) (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k))^T \right] \\ = & \mathbb{E}^{(i)} \left[\left[(\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k}^{(i)})) + (\mathbf{h}(\hat{\mathbf{x}}_{k|k}^{(i)}) - \mathbf{h}(\mathbf{x}_k)) \right] \cdot [\cdot]^T \right] \\ = & (\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k}^{(i)})) (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_{k|k}^{(i)}))^T + \mathbf{P}_{zz}^{(i)} \\ = & (\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k}^{(i)})) (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_{k|k}^{(i)}))^T + \frac{1}{2n} \\ & \cdot \sum_{j=1}^{2n} \left(\mathbf{h}(\mathbf{x}_{j|k}^{(i)}) - \hat{\mathbf{z}}_{k|k}^{(i)} \right) \left(\mathbf{h}(\mathbf{x}_{j|k}^{(i)}) - \mathbf{z}_{k|k}^{(i)} \right)^T \end{aligned} \quad (33)$$

The i -th cubature point $\mathbf{x}_{j|k}^{(i)}$ ($j = 1, 2, \dots, 2n$) and the predicted measurement $\hat{\mathbf{z}}_{k|k}^{(i)}$ are computed as follows:

$$\begin{aligned} \mathbf{x}_{j|k}^{(i)} = & \hat{\mathbf{x}}_{k|k}^{(i)} + \mathbf{S}_{k|k}^{(i)} \boldsymbol{\xi}_j, j = 1, 2, \dots, 2n \\ \hat{\mathbf{z}}_{k|k}^{(i)} = & \frac{1}{2n} \cdot \sum_{j=1}^{2n} \mathbf{h}(\mathbf{x}_{j|k}^{(i)}) \end{aligned} \quad (34)$$

where $\mathbf{S}_{k|k}^{(i)}$ is the Cholesky decomposition of $\mathbf{P}_{k|k}^{(i)}$ satisfying $\mathbf{P}_{k|k}^{(i)} = \mathbf{S}_{k|k}^{(i)} \cdot (\mathbf{S}_{k|k}^{(i)})^T$.

Contrasting (13) with (32), the approximate posterior PDF $q^{(i+1)}(\eta_k)$ can be given by

$$q^{(i+1)}(\eta_k) = \mathcal{G}\left(\eta_k; \omega_k^{(i+1)}, \zeta_k^{(i+1)}\right) \quad (35)$$

where $\omega_k^{(i+1)} = m + v/2$, $\zeta_k^{(i+1)} = (v + \text{tr}(\mathcal{R}_k^{(i)} \mathbf{R}_k^{-1}))/2$

Let $\phi = \mathcal{T}_k$ and substituting (28) in (25) results in

$$\begin{aligned} \ln q^{(i+1)}(\mathcal{T}_k) = & -\frac{n + t_k + 2}{2} \ln |\mathcal{T}_k| \\ & - \frac{\text{tr}\left(\left(\mathbf{T}_k + \mathbf{E}^{(i+1)}[\zeta_k] \mathcal{P}_k^{(i)}\right) \mathcal{T}_k^{-1}\right)}{2} + c_{\mathcal{T}} \end{aligned} \quad (36)$$

Contrasting (36) with (18), the approximate posterior PDF $q^{(i+1)}(\mathcal{T}_k)$ can be given by

$$q^{(i+1)}(\mathcal{T}_k) = \text{IW}\left(\mathcal{T}_k; t_k^{(i+1)}, \mathbf{T}_k^{(i+1)}\right) \quad (37)$$

where $t_k^{(i+1)} = t_k + 1$, $\mathbf{T}_k^{(i+1)} = \mathbf{T}_k + \mathbf{E}^{(i+1)}[\zeta_k] \mathcal{P}_k^{(i)}$, and $\mathbf{E}^{(i+1)}[\zeta_k] = \alpha_k^{(i+1)} / \beta_k^{(i+1)}$.

Let $\phi = \mathbf{x}_k$ and substituting (28) into (25) results in

$$\begin{aligned} \ln q^{(i+1)}(\mathbf{x}_k) = & -\frac{\mathbf{E}^{(i+1)}[\eta_k] (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k))^T \mathbf{R}_k^{-1} (\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k))}{2} \\ & - \frac{\mathbf{E}^{(i+1)}[\zeta_k] (\mathbf{x}_k - \mathbf{x}_{k|k-1})^T \mathbf{E}^{(i+1)}[\mathcal{T}_k^{-1}] (\mathbf{x}_k - \mathbf{x}_{k|k-1})}{2} \\ & + c_{\mathbf{x}} \end{aligned} \quad (38)$$

where $\mathbf{E}^{(i+1)}[\eta_k] = \omega_k^{(i+1)} / \zeta_k^{(i+1)}$, $\mathbf{E}^{(i+1)}[\mathcal{T}_k^{-1}] = (t_k^{(i+1)} - n - 1) (\mathbf{T}_k^{(i+1)})^{-1}$.

According to the Bayesian theory, the posterior PDF can be approximated by

$$q^{(i+1)}(\mathbf{x}_k) = \frac{p^{(i+1)}(\mathbf{z}_k | \mathbf{x}_k) p^{(i+1)}(\mathbf{x}_k | \mathbf{Z}_{k-1})}{\int p^{(i+1)}(\mathbf{z}_k | \mathbf{x}_k) p^{(i+1)}(\mathbf{x}_k | \mathbf{Z}_{k-1}) d\mathbf{x}_k} \quad (39)$$

Exploiting (38), $p^{(i+1)}(\mathbf{z}_k | \mathbf{x}_k)$ and $p^{(i+1)}(\mathbf{x}_k | \mathbf{Z}_{k-1})$ can be written as

$$\begin{aligned} p^{(i+1)}(\mathbf{z}_k | \mathbf{x}_k) &= \mathbf{N}\left(\mathbf{z}_k; \mathbf{h}(\mathbf{x}_k), \bar{\mathbf{R}}_k^{(i+1)}\right) \\ p^{(i+1)}(\mathbf{x}_k | \mathbf{Z}_{k-1}) &= \mathbf{N}\left(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right) \end{aligned} \quad (40)$$

where $\bar{\mathbf{R}}_k^{(i+1)} = \mathbf{R}_k / \mathbf{E}^{(i+1)}[\eta_k]$, $\bar{\mathbf{P}}_{k|k-1}^{(i+1)} = (\mathbf{E}^{(i+1)}[\mathcal{T}_k^{-1}])^{-1} / \mathbf{E}^{(i+1)}[\zeta_k]$.

Since Gaussian distribution can approximate the posterior PDF more accurate than the Student's t distribution [24], the approximate posterior PDF $q^{(i+1)}(\mathbf{x}_k)$ can be formulated by

$$q^{(i+1)}(\mathbf{x}_k) = \mathbf{N}\left(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{(i+1)}, \mathbf{P}_{k|k}^{(i+1)}\right) \quad (41)$$

According to the basic equations of information filter, we have

$$\begin{aligned} \bar{\mathbf{Y}}_{k|k-1}^{(i+1)} &= \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-1} = \mathbf{E}^{(i+1)}[\zeta_k] \left(\mathbf{E}^{(i+1)}[\mathcal{T}_k^{-1}]\right) \\ \mathbf{y}_{k|k-1}^{(i+1)} &= \bar{\mathbf{Y}}_{k|k-1}^{(i+1)} \hat{\mathbf{x}}_{k|k-1} \end{aligned} \quad (42)$$

Based on the linear error propagation, the cross covariance $\bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)}$ is approximated by

$$\bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)} \approx \bar{\mathbf{P}}_{k|k-1}^{(i+1)} \mathbf{H}_k^{(i+1)} \quad (43)$$

Multiplying $[(\bar{\mathbf{P}}_{k|k-1}^{(i+1)})^{-1} \bar{\mathbf{P}}_{k|k-1}^{(i+1)}]$ and $[(\bar{\mathbf{P}}_{k|k-1}^{(i+1)})^{-1} \bar{\mathbf{P}}_{k|k-1}^{(i+1)}]^T$ on the right hand side of (8) and (9) yields

$$\begin{aligned} \mathbf{i}_k^{(i+1)} &= \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-1} \bar{\mathbf{P}}_{k|k-1}^{(i+1)} \left(\mathbf{H}_k^{(i+1)}\right)^T \mathbf{R}_k^{-1} \\ &\cdot \left[\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}^{(i+1)} + \mathbf{H}_k^{(i+1)} \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^T \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-T} \hat{\mathbf{x}}_{k|k-1} \right] \end{aligned} \quad (44)$$

$$\mathbf{I}_k^{(i+1)} = \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-1} \bar{\mathbf{P}}_{k|k-1}^{(i+1)} \mathbf{H}_k^T \left(\bar{\mathbf{R}}_k^{(i+1)}\right)^{-1} \mathbf{H}_k \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^T \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-T} \quad (45)$$

Using (43) to simplify (44) and (45) results in

$$\begin{aligned} \mathbf{i}_k^{(i+1)} &= \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-1} \bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)} \mathbf{R}_k^{-1} \\ &\cdot \left[\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}^{(i+1)} + \left(\bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)}\right)^T \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-T} \hat{\mathbf{x}}_{k|k-1} \right] \end{aligned} \quad (46)$$

$$\mathbf{I}_k^{(i+1)} = \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-1} \bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)} \left(\bar{\mathbf{R}}_k^{(i+1)}\right)^{-1} \left(\bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)}\right)^T \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)}\right)^{-T} \quad (47)$$

The predicted measurement $\hat{\mathbf{z}}_{k|k-1}^{(i+1)}$ and cross covariance $\bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)}$ are computed as follows:

$$\begin{aligned} \hat{\mathbf{z}}_{k|k-1}^{(i+1)} &= \frac{1}{2n} \cdot \sum_{j=1}^{2n} \mathbf{h}\left(\boldsymbol{\chi}_{j,k|k-1}^{(i+1)}\right) \\ \bar{\mathbf{P}}_{xz,k|k-1}^{(i+1)} &= \frac{1}{2n} \cdot \sum_{j=1}^{2n} \left(\boldsymbol{\chi}_{j,k|k-1}^{(i+1)} - \hat{\mathbf{x}}_{k|k-1}\right) \left(\mathbf{h}\left(\boldsymbol{\chi}_{j,k|k-1}^{(i+1)}\right) - \mathbf{z}_{k|k-1}^{(i+1)}\right)^T \end{aligned} \quad (48)$$

FIGURE 1: Diagram of the proposed NRSTCIF.

where $\chi_{j,k|k-1}^{(i+1)}$ ($i = 1, 2, \dots, 2n$) is the i -th cubature point generated by

$$\chi_{j,k|k-1}^{(i+1)} = \hat{\mathbf{x}}_{k|k-1} + \bar{\mathbf{S}}_{k|k-1}^{(i+1)} \boldsymbol{\xi}_j, j = 1, 2, \dots, 2n \quad (49)$$

and $\bar{\mathbf{S}}_{k|k-1}^{(i+1)}$ is the Cholesky decomposition of $\bar{\mathbf{P}}_{k|k-1}^{(i+1)}$.

Based on the information contributions given by (46) and (47), the information state and matrix are updated as

$$\begin{aligned} \hat{\mathbf{y}}_{k|k}^{(i+1)} &= \mathbf{y}_{k|k-1}^{(i+1)} + \mathbf{i}_k^{(i+1)} \\ \mathbf{Y}_{k|k}^{(i+1)} &= \bar{\mathbf{Y}}_{k|k-1}^{(i+1)} + \mathbf{I}_k^{(i+1)} \end{aligned} \quad (50)$$

Then, the estimated state and covariance can be recovered as

$$\mathbf{P}_{k|k}^{(i+1)} = \left(\mathbf{Y}_{k|k}^{(i+1)} \right)^{-1} \quad (51)$$

$$\hat{\mathbf{x}}_{k|k}^{(i+1)} = \mathbf{P}_{k|k}^{(i+1)} \hat{\mathbf{y}}_{k|k}^{(i+1)}$$

After the iterations, the approximate PDFs given in (23) are updated as

$$\begin{aligned} q(\mathbf{x}_k) &\approx \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^{(d)}, \mathbf{P}_{k|k}^{(d)}) \\ q(\varsigma_k) &\approx \mathcal{G}(\varsigma_k; \alpha_{k|k}^{(d)}, \beta_{k|k}^{(d)}) \\ q(\eta_k) &\approx \mathcal{G}(\eta_k; \bar{\omega}_{k|k}^{(d)}, \zeta_{k|k}^{(d)}) \\ q(\mathcal{T}_k) &= \text{IW}(\mathcal{T}_k; t_k^{(d)}, \mathbf{T}_k^{(d)}) \end{aligned} \quad (52)$$

where d is the number of iterations (NOI).

FIGURE 2: Tracking performance of four filters; (a) Position estimation error, (b) Velocity estimation error, (c) Turn rate estimation error.

The main advantage of the IF is its ability to fuse multisource measurements simply by adding the information contributions to the information matrix and state. Suppose N sensors is available, the measurement equations can be given by

$$\mathbf{z}_{\kappa,k} = \mathbf{h}_{\kappa}(\mathbf{x}_k) + \mathbf{v}_{\kappa,k}, \kappa = 1, 2, \dots, N \quad (53)$$

where the subscript κ denotes the κ -th sensor, $\mathbf{v}_{\kappa,k} \sim \text{St}(\mathbf{v}_{\kappa,k}; \mathbf{0}, \mathbf{R}_{\kappa,k}, \nu)$ is the measurement noise.

Then, the information matrix and information state can be updated by

$$\begin{aligned} \hat{\mathbf{y}}_{k|k}^{(i+1)} &= \mathbf{y}_{k|k-1}^{(i+1)} + \sum_{\kappa=1}^N \mathbf{i}_{\kappa,k}^{(i+1)} \\ \mathbf{Y}_{k|k}^{(i+1)} &= \bar{\mathbf{Y}}_{k|k-1}^{(i+1)} + \sum_{\kappa=1}^N \mathbf{I}_{\kappa,k}^{(i+1)} \end{aligned} \quad (54)$$

where the information contributions of κ -th sensor are computed by

$$\begin{aligned}
\mathbf{i}_{\kappa,k}^{(i+1)} &= \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)} \right)^{-1} \bar{\mathbf{P}}_{\kappa,xz,k|k-1}^{(i+1)} \bar{\mathbf{R}}_k^{-1} \left[\mathbf{z}_{\kappa,k} - \hat{\mathbf{z}}_{\kappa,k|k-1}^{(i+1)} \right. \\
&\quad \left. + \left(\bar{\mathbf{P}}_{\kappa,xz,k|k-1}^{(i+1)} \right)^T \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)} \right)^{-T} \hat{\mathbf{x}}_{k|k-1} \right], \kappa = 1, 2, \dots, N \\
\mathbf{I}_{\kappa,k}^{(i+1)} &= \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)} \right)^{-1} \bar{\mathbf{P}}_{\kappa,xz,k|k-1}^{(i+1)} \left(\bar{\mathbf{R}}_{\kappa,k}^{(i+1)} \right)^{-1} \\
&\quad \cdot \left(\bar{\mathbf{P}}_{\kappa,xz,k|k-1}^{(i+1)} \right)^T \left(\bar{\mathbf{P}}_{k|k-1}^{(i+1)} \right)^{-T}, \kappa = 1, 2, \dots, N
\end{aligned} \tag{55}$$

To better illustrate the proposed filtering algorithm, the computational procedures are summarized in Figure 1.

4. Simulation

To demonstrate the feasibility of the proposed algorithm, a target tracking problem is investigated in this section. The target moves in a plane with an unknown speed and unknown constant turn rate, it is observed by multiple radars in clutter. Due to the rapid motion and clutter, the heavy-tailed noises are introduced.

The state equation is given by

$$\mathbf{x}_k = \begin{bmatrix} 1 & \frac{\sin(\omega_{k-1}T)}{\omega_{k-1}} & 0 & \frac{\cos(\omega_{k-1}T) - 1}{\omega_{k-1}} & 0 \\ 0 & \cos(\omega_{k-1}T) & 0 & -\sin(\omega_{k-1}T) & 0 \\ 0 & \frac{1 - \cos(\omega_{k-1}T)}{\omega_{k-1}} & 1 & \frac{\sin(\omega_{k-1}T)}{\omega_{k-1}} & 0 \\ 0 & \sin(\omega_{k-1}T) & 0 & \cos(\omega_{k-1}T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_{k-1} \tag{56}$$

where the state vector $\mathbf{x}_k$ is defined as $\mathbf{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k, \omega_k]^T$, ω_k is the turn rate, $T = 1$ s is the sampling interval, and $\mathbf{w}_{k-1}$ is the process noise generated by

$$\mathbf{w}_{k-1} = \begin{cases} \mathbf{N}(\mathbf{0}, \mathbf{Q}_{k-1}) & \text{with probability } 1 - p \\ \mathbf{N}(\mathbf{0}, 50\mathbf{Q}_{k-1}) & \text{with probability } p \end{cases} \tag{57}$$

where $p = 0.05$ is probability of outlier, $\mathbf{Q}_{k-1}$ is given by

$$\mathbf{Q}_{k-1} = \begin{bmatrix} T^3/3 & T^2/2 & 0 & 0 & 0 \\ T^2/2 & T & 0 & 0 & 0 \\ 0 & 0 & T^3/3 & T^2/2 & 0 \\ 0 & 0 & T^2/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.75 \times 10^{-3} T \end{bmatrix} \tag{58}$$

TABLE 1: Computational time comparisons.

Filter	CIF	HCIF	STCIF	NRSTCIF
Time(s)	0.11	0.19	0.24	0.5

FIGURE 3: Averaged RMSEs with different NOIs.

The measurement equation of j -th radar is given by

$$\mathbf{z}_{j,k} = \begin{bmatrix} \sqrt{(x_k - x_{j,s})^2 + (y_k - y_{j,s})^2} \\ \text{atan2}(y_k - y_{j,s}, x_k - x_{j,s}) \end{bmatrix} + \mathbf{v}_k \tag{59}$$

where $(x_{j,s}, y_{j,s})$ is the position of j -th radar, $\mathbf{v}_k$ is the measurement noise generated by

$$\mathbf{v}_k = \begin{cases} \mathbf{N}(\mathbf{0}, \mathbf{R}_k) & \text{with probability } 1 - p \\ \mathbf{N}(\mathbf{0}, 100\mathbf{R}_k) & \text{with probability } p \end{cases} \tag{60}$$

and $\mathbf{R}_k$ is given by

$$\mathbf{R}_k = \begin{bmatrix} 100\text{m}^2 & 0 \\ 0 & 100\text{mrad}^2 \end{bmatrix} \tag{61}$$

The locations of the four sensors are given as

$$\begin{cases} x_{1,s} = -30000\text{m} \\ y_{1,s} = -5000\text{m} \end{cases} \quad \begin{cases} x_{2,s} = 0\text{m} \\ y_{2,s} = 35000\text{m} \end{cases} \tag{62}$$

$$\begin{cases} x_{3,s} = 30000\text{m} \\ y_{3,s} = -5000\text{m} \end{cases} \quad \begin{cases} x_{4,s} = 0\text{m} \\ y_{4,s} = -40000\text{m} \end{cases}$$

FIGURE 4: Performance comparison of three filters with different probabilities of outlier. (a) Position estimation error. (b) Velocity estimation error. (c) Turn rate estimation error.

The initial state estimation $\hat{\mathbf{x}}_{0|0}$ is generated randomly based on the true initial value $\mathbf{x}_0$ and initial covariance $\mathbf{P}_0$ given by

$$\mathbf{x}_0 = [1000\text{m}, 300\text{m/s}, 1000\text{m}, 0, -3^\circ/\text{s}]^T$$

$$\mathbf{P}_0 = \text{diag} [100\text{m}^2, 10\text{m}^2/\text{s}^2, 100\text{m}^2, 10\text{m}^2/\text{s}^2, 100\text{mrad}^2/\text{s}^2]^T \quad (63)$$

In the simulation, the dof parameter, tuning parameter, and NOI of the proposed filter are set as $\mu = \nu = 5$, $\tau = 5$, and $d = 10$, respectively. The root mean square error (RMSE) is chosen as the performance metric. Based on 1000 indepen-

dent Monte Carlo simulation runs, Figure 2 compares the performance of cubature information filter (CIF) [17], Huber-based cubature information filter (HCIF) [19], Student's t -based cubature information filter (STCIF) [23], and the proposed filter. It can be seen that the CIF has the biggest RMSEs among four filters. This is not surprise, since the CIF is essentially a minimizer of a weighted least-squares criterion, which is sensitive to heavy-tailed noises. In addition, one can also observe that the RMSEs of the proposed filter are smaller than the existing HCIF and STCIF, which indicates the superiority of the proposed filter.

The average computational costs of the various filters for each Monte-Carlo run are given in Table 1. It can be seen that

the computational time of the NRSTCIF is the longest, and that of the CIF is the shortest. It takes the NRSTCIF about 5 times as long as the CIF. Therefore, the algorithm proposed in this paper does not increase the computation time dramatically.

To clarify the relation between estimation accuracy and NOI, another 1000 independent Monte Carlo simulation runs with different NOIs are conducted based on the proposed filter, and the simulation results are summarized in Figure 3. It can be seen that the averaged RMSEs decrease as the NOI increase when $\text{NOI} = 1, 2$, and 3, but the averaged RMSEs are almost same when $\text{NOI} \geq 4$. Therefore, it is unnecessary to choose the NOI as big as possible, and the result is an important reference for NOI choosing.

Figure 4 shows the RMSEs of the CIF, HCIF, and the proposed filter with different probabilities of outlier $p = 0.01, 0.02, \dots, 0.1$. It can be seen the proposed filter is superior to the CIF and HCIF for all values of p . And the averaged RMSEs of CIF and HCIF increase more rapidly than the proposed NRSTCIF as the increase of p , which indicate the superiority of the NRSTCIF is more and more significant. In other words, the proposed NRSTCIF is more promising when the probability of outlier is larger.

5. Conclusions

In this paper, a novel robust Student's t -based cubature information filter is proposed to handle the heavy-tailed process and measurement noises in nonlinear multisensor estimation. Based on the cubature rule, the one-step predicted state and the nominal covariance are propagated. To get the approximate solution of the posterior PDF, the predictive PDF and the likelihood PDF described by Student's t distribution are transformed into a hierarchical Gaussian form. To approximate the joint posterior PDF, the VB approximation method is utilized. After fixed point iterations, the state, scale matrix, and posterior densities are estimated. And the simulation results demonstrate the superiority of the proposed filter.

Data Availability

There is no underlying data related to the submission.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- [1] X. Wang, Q. Zhang, and H. Li, "An autonomous navigation scheme based on starlight, geomagnetic and gyros with information fusion for small satellites," *Acta Astronautica*, vol. 94, no. 2, pp. 708–717, 2014.
- [2] Y. Beaudoin, A. Desbiens, E. Gagnon, and R. Landry Jr., "Observability of satellite launcher navigation with INS, GPS, attitude sensors and reference trajectory," *Acta Astronautica*, vol. 142, pp. 277–288, 2018.
- [3] D. Willner, C. B. Chang, and K. P. Dunn, "Kalman filter algorithms for a multi-sensor system," in *1976 IEEE Conference on Decision and Control including the 15th Symposium on Adaptive Processes*, pp. 570–574, Clearwater, FL, USA, 1976.
- [4] Q. Gan and C. J. Harris, "Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 1, pp. 273–279, 2001.
- [5] Y. Wang, W. Zheng, S. Sun, and L. Li, "Robust information filter based on maximum correntropy criterion," *Journal of Guidance, Control, and Dynamics*, vol. 39, no. 5, pp. 1126–1131, 2016.
- [6] M. Roth, E. Özkan, and F. Gustafsson, "A Student's T Filter for Heavy Tailed Process and Measurement Noise," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5770–5774, Vancouver, BC, Canada, 2013.
- [7] A. Jazwinski, *Stochastic Processing and Filtering Theory*, Academic Press, New York, USA, 1970.
- [8] S. Julier, J. Uhlmann, and H. F. Durrantwhyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.
- [9] H. Leeghim, Y. Choi, and B. A. Jaroux, "Uncorrelated unscented filtering for spacecraft attitude determination," *Acta Astronautica*, vol. 67, no. 1-2, pp. 135–144, 2010.
- [10] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 910–927, 2000.
- [11] M. Nørgaard, N. K. Poulsen, and O. Ravn, "New developments in state estimation for nonlinear systems," *Automatica*, vol. 36, no. 11, pp. 1627–1638, 2000.
- [12] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.
- [13] H. Renke, S. Chen, H. Wu, Z. Liu, and J. Chen, "Optimal Maneuver Strategy of Observer for Bearing-Only Tracking in Threat Environment," *International Journal of Aerospace Engineering*, vol. 2018, Article ID 7901917, 9 pages, 2018.
- [14] B. Jia and M. Xin, "Multiple sensor estimation using a new fifth-degree cubature information filter," *Transactions of the Institute of Measurement and Control*, vol. 37, no. 1, pp. 15–24, 2014.
- [15] S. E. Webster, J. M. Walls, L. L. Whitcomb, and R. M. Eustice, "Decentralized extended information filter for single-Beacon cooperative acoustic navigation: theory and experiments," *IEEE Transactions on Robotics*, vol. 29, no. 4, pp. 957–974, 2013.
- [16] D. J. Lee, "Nonlinear estimation and multiple sensor fusion using unscented information filtering," *IEEE Signal Processing Letters*, vol. 15, pp. 861–864, 2008.
- [17] K. P. B. Chandra, D. W. Gu, and I. Postlethwaite, "Square root cubature information filter," *IEEE Sensors Journal*, vol. 13, no. 2, pp. 750–758, 2013.
- [18] X. Wang, D. Gong, L. Xu, X. Shao, and D. Duan, "Laser radar based relative navigation using improved adaptive Huber filter," *Acta Astronautica*, vol. 68, no. 11-12, pp. 1872–1880, 2011.
- [19] X. Wang, N. Cui, and J. Guo, "Huber-based unscented filtering and its application to vision-based relative navigation," *IET Radar, Sonar and Navigation*, vol. 4, no. 1, pp. 134–141, 2010.

- [20] X. Wang, W. Qin, N. Cui, and Y. Wang, "Robust high-degree cubature information filter and its application to trajectory estimation for ballistic missile," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 232, no. 12, pp. 2364–2377, 2017.
- [21] 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.
- [22] G. Wang, N. Li, and Y. Zhang, "Maximum correntropy unscented Kalman and information filters for non-Gaussian measurement noise," *Journal of the Franklin Institute*, vol. 354, no. 18, pp. 8659–8677, 2017.
- [23] Y. Huang, Y. Zhang, N. Li, and J. Chambers, "Robust student's t based nonlinear filter and smoother," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2586–2596, 2016.
- [24] Y. Huang, Y. Zhang, Z. Wu, N. Li, and J. Chambers, "A novel robust t-Based Kalman filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 3, pp. 1545–1554, 2017.
- [25] Y. Huang, Y. Zhang, B. Xu, and J. Chambers, "A new outlier-robust Student's t based Gaussian approximate filter for cooperative localization," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 5, pp. 2380–2386, 2017.
- [26] Y. Huang, Y. Zhang, and J. A. Chambers, "A novel Kullback–Leibler divergence minimization-based adaptive Student's t-filter," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5417–5432, 2019.
- [27] Y. Huang, Y. Zhang, Y. Zhao, and J. A. Chambers, "A novel robust Gaussian–Student's t Mixture distribution based Kalman filter," *IEEE Transactions on Signal Processing*, vol. 67, no. 13, pp. 3606–3620, 2019.
- [28] Y. Huang, Y. Zhang, Z. Wu, N. Li, and J. Chambers, "A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 594–601, 2018.
- [29] Y. Huang, Y. Zhang, P. Shi, Z. Wu, J. Qian, and J. A. Chambers, "Robust Kalman filters based on Gaussian scale mixture distributions with application to target tracking," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 10, pp. 2082–2096, 2019.
- [30] P. Dong, Z. Jing, H. Leung, and K. Shen, "Variational Bayesian Adaptive Cubature Information Filter Based on Wishart Distribution," *IEEE Trans. Automat. Contr.*, vol. 62, no. 11, pp. 6051–6057, 2017.