sensors

MDPI

Article

An Adaptive Filter for Nonlinear Multi-Sensor Systems with Heavy-Tailed Noise

Xiangxiang Dong [1,2,3], Luigi Chisci [4] and Yunze Cai [1,2,3,*]

1 Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China; js.danesir@sjtu.edu.cn
2 Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China
3 Key Laboratory of Marine Intelligent Equipment and System of Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China
4 Department of Information Engineering, Università di Firenze, 50139 Firenze, Italy; luigi.chisci@unifi.it
* Correspondence: yzcai@sjtu.edu.cn

Received: 25 August 2020; Accepted: 23 November 2020; Published: 26 November 2020

Abstract: Aiming towards state estimation and information fusion for nonlinear systems with heavy-tailed measurement noise, a variational Bayesian Student's t-based cubature information filter (VBST-CIF) is designed. Furthermore, a multi-sensor variational Bayesian Student's t-based cubature information feedback fusion (VBST-CIFF) algorithm is also derived. In the proposed VBST-CIF, the spherical-radial cubature (SRC) rule is embedded into the variational Bayes (VB) method for a joint estimation of states and scale matrix, degree-of-freedom (DOF) parameter, as well as an auxiliary parameter in the nonlinear system with heavy-tailed noise. The designed VBST-CIF facilitates multi-sensor fusion, allowing to derive a VBST-CIFF algorithm based on multi-sensor information feedback fusion. The performance of the proposed algorithms is assessed in target tracking scenarios. Simulation results demonstrate that the proposed VBST-CIF/VBST-CIFF outperform the conventional cubature information filter (CIF) and cubature information feedback fusion (CIFF) algorithms.

Keywords: nonlinear multi-sensor system; heavy-tailed noise; student's t distribution; spherical-radial cubature rule; information fusion

1. Introduction

The Kalman filter (KF) is an optimal state estimator for linear state-space systems [1–3]. It is widely used, owing to its optimality, in many applications like, e.g., localization, control, target tracking, and signal processing [4–17]. In reality, however, systems are usually characterized by strong non-linearities which make the conventional KF inappropriate. To this end, nonlinear filtering methods have been developed like, e.g., function approximation, deterministic sampling, and Monte Carlo estimation methods [18–20]. The function approximation method adopted by the extended Kalman filter (EKF) approximates the nonlinear system equations through truncated Taylor expansions [21]. However, the Jacobian matrix is not computable for systems with non-smooth non-linearities [21]. As for the deterministic sampling method, its main representatives are the unscented Kalman filter (UKF), cubature Kalman filter (CKF), etc. [22,23]. Unfortunately, the state error covariance matrix (SECM) in UKF may result non-positive definite for high-dimensional systems, possibly leading to filter divergence [23]. To overcome the drawbacks of UKF, Arasaratnam and Haykin proposed CKF in 2009 [23]. CKF approximates the posterior probability density function (PDF) by means of the spherical-radial cubature (SRC) rule, thus resulting in improved filtering accuracy for high-dimensional systems [24]. Moreover, compared to the Monte Carlo estimation approach exploited by particle filters (PFs), CKF is not affected by particle depletion issues.

Recall that the KF (and also EKF, UKF, and CKF) can also be implemented in the alternative information filter (IF) form wherein the inverse covariance (information) matrix is propagated instead of the covariance. In this respect, the IF form greatly simplifies the measurement update compared to the traditional covariance form [25]. In this way, it facilitates multi-sensor fusion, wherein measurements from different sensors are fused via directly adding information contributions to the information matrix and vector [26]. Considering the above-stated superiority in the multi-sensor case, the cubature information filter (CIF) will be adopted as a filtering approach in this paper.

Unfortunately, the existing CIF assumes the known and constant measurement noise covariance matrix (MNCM) [27]. However, due to the complex environment where the sensor is located, it is prone to measurement outliers. This might induce a heavy-tailed measurement noise [28]. Actually, an outlier is characterized by the abnormality of its measurement value; correspondingly, outliers might give rise to a heavy-tailed distribution of the measurement noise. There are many practical applications where there is heavy-tailed noise [29–32]. For instance, in vision-aided inertial navigation, either computer vision data contaminated by outliers or sonar data corrupted by phase noise may lead to heavy-tailed noise [29]. In the detection of FM signals, heavy-tailed noise can be produced by thunderstorms or an iceberg breakup in under-ice acoustics [30]. In target tracking scenarios, measurement outliers come from unreliable sensors and/or targets [31], while in target detection with synthetic aperture radar, heavy-tailed noise occurs due to clutter and buried or obscured observed targets [32], etc. Compared to Gaussian noise (hypothesized by conventional CIF), heavy-tailed noise is characterized by greater uncertainty in its distribution tails, yielding the heavy-tailed shape of the PDF [33]. Under the above circumstance, the filtering accuracy of CIF may deteriorate and filter divergence may occur, due to the discrepancy between the assumed noise statistics and the true one.

To tackle the filtering problem under heavy-tailed noise, several adaptive methods have been developed, including generalized maximum likelihood [34], multi-model [35], and joint estimation [36] methods. Specifically, a representative of generalized maximum likelihood estimation is the Huber-based filter, which minimizes the combined l_1 and l_2 norms [37] and, however, leads to limited estimation performance [38]. An alternative is the multi-model approach, which models the unknown noise as a random parameter switching according to a Markov Chain and is able to adaptively estimate the system state [35]. Unfortunately, it is limited by the selected sets of models. The greater is the deviation between the selected models and real noise model, the worse the estimation performance of this method [36]. Furthermore, its computational complexity will increase significantly as the number of unknown parameters and selected models increases. The joint estimation method includes expectation-maximization (EM) and variational Bayes (VB) approaches [36]. The EM approach identifies and estimates system states as well as hidden variables through expectation (E) and maximization (M) steps. However, it is difficult to obtain the analytical solution for E and M estimates in the case of high-dimensional systems.

Conversely, another joint estimation method, the VB approach, can avoid the aforementioned drawbacks of the EM approach for high-dimensional systems. By searching an approximation for the true joint distribution, the VB approach is widely used in the adaptive joint estimation of noise statistics [27,29,38–42]. Agamennoni et al. proposed a robust Kalman filter by exploiting a flexible model [39] and also introduced a structured variational approximation approach for heavy-tailed noise [40]. However, the use of these filters is restricted to linear systems. Huang et al. [38] also presented a linear Student's t (ST) filter. By modeling heavy-tailed noise with ST distribution, the system state as well as noise statistics are adaptively estimated through VB iterations. Unfortunately, this approach is only applicable to linear systems, too. Although there has been some research work on nonlinear systems in [29,41,42], the proposed filters rely on the covariance filter form, instead of the IF form, in order to jointly estimate system state, SECM, and noise statistics, thus resulting into complex computation. Moreover, the current VB approach cannot be embedded into nonlinear CIF directly, since the computation of noise parameters like the scale matrix involves a nonlinear function and, therefore, requires the use of the SRC rule. Hence, to jointly estimate states and parameters of the noise

statistics (especially the scale matrix) via the information filter framework is still an open problem in nonlinear systems with heavy-tailed noise. In addition, the existing research work is only able to deal with single-sensor estimation problems. For nonlinear multi-sensor systems with heavy-tailed noise, the current state-of-art does not allow to estimate and fuse the states as well as noise statistics.

To conclude, there still exist the following issues to be addressed for the state estimation of nonlinear multi-sensor systems. On the one hand, the nonlinear information filter cannot be embedded into the current VB framework, since the scale matrix cannot be directly computed in the same way as for linear or nonlinear systems in the covariance filter form. On the other hand, it is not suited to multi-sensor systems. There are different heavy-tailed noise signals when multiple sensors are working in a multi-sensor system. This makes the conventional filter for the single-sensor case fail to fuse different estimated states. Thus, for nonlinear systems with heavy-tailed noise, especially systems with multiple sensors, it is still an open issue how to jointly estimate system state and noise statistics by combining information from different sensors.

To solve the above mentioned problems, an adaptive variational Bayesian Student's t-based cubature information filter (VBST-CIF) algorithm is developed. Furthermore a variational Bayesian Student's t-based cubature information feedback fusion (VBST-CIFF) algorithm is proposed for nonlinear multi-sensor systems with heavy-tailed noise. Key contributions of this paper are the following.

(1) The information filter form is adopted for simplified computation and facilitation of multi-sensor fusion.
(2) A novel VBST-CIF algorithm for nonlinear systems with heavy-tailed noise is proposed. The SRC rule is introduced into the VB approach for joint estimation of states and noise statistics, by employing the ST distribution for modeling heavy-tailed noise.
(3) The proposed VBST-CIF algorithm is further extended to multi-sensor fusion, deriving a novel VBST-CIFF algorithm. The proposed VBST-CIFF algorithm facilitates multi-sensor fusion in nonlinear systems with different heavy-tailed measurement noise statistics for each sensor.
(4) Simulation and experimental results show that the proposed VBST-CIF/VBST-CIFF algorithms outperform conventional CIF and cubature information feedback fusion (CIFF) algorithms in scenarios concerning nonlinear systems with heavy-tailed noise.

The rest of the paper is organized as follows. First, the problem of state estimation for nonlinear systems with heavy-tailed measurement noise is formulated in Section 2. Then, the SRC rule and ST distribution are introduced in Section 3, wherein the ST distribution is utilized to model heavy-tailed noise and the VBST-CIF algorithm is also derived. Furthermore, a VBST-CIFF algorithm is provided for multi-sensor fusion of nonlinear systems with heavy-tailed noise in Section 4. In Section 5, the proposed VBST-CIF and VBST-CIFF algorithms are tested in nonlinear target tracking scenarios. Conclusions are provided in Section 6.

2. Problem Formulation

Consider the following model:

$$x_k = f(x_{k-1}) + w_{k-1} \tag{1}$$

$$y_k = h(x_k) + v_k \tag{2}$$

where: k denotes the discrete time index; x_k and y_k are the state and measurement vector, respectively; f and h denote state transition and, respectively, measurement function; w_k is a zero-mean Gaussian process noise with covariance Q_k, i.e., $w_k \sim \mathcal{N}(0, Q_k)$, while v_k is heavy-tailed measurement noise; $\mathcal{N}(\mu, \Sigma)$ denotes a Gaussian random variable with mean μ and variance Σ. Moreover, it is assumed that w_i and v_j are uncorrelated for any i and j.

Conventional CIF can deal with state estimation in the Gaussian case. It has been introduced in [43] under the assumption that measurement noise has Gaussian distribution. Unfortunately,

for systems (1) and (2) affected by heavy-tailed measurement noise, it may lead to non-accurate estimation because of the deviation between the assumed noise model and the actual one.

To this end, Huang et al. [38] have proposed an adaptive filter for linear systems. Nevertheless, such a filter is limited to linear single-sensor systems and is not suited to multi-sensor nonlinear systems. Although there are some methods for nonlinear state estimation, how to calculate the matrix parameter of heavy-tailed noise is still an open issue. Specifically, the posterior distribution parameters of the heavy-tailed noise contain a matrix related to the measurement matrix in traditional linear systems. Unfortunately, such a matrix cannot be directly obtained for nonlinear systems, thus implying that the analytical solution for joint estimation of system state as well as the noise parameters is not possible. On the other hand, for multi-sensor systems, multiple noise matrices need to be estimated, one for each sensor s. Hence, how to estimate these parameters together with the system state and fuse them still remains another open issue.

To conclude, the following problems still need to be addressed for the state estimation of a nonlinear system with heavy-tailed measurement noise.

(1) The conventional CIF assumes Gaussian measurement noise distribution. For systems with heavy-tailed noise, CIF is not able to estimate states and the noise statistics simultaneously.
(2) The current VB approach based on the conventional KF framework is not suited to nonlinear systems. In particular, the unknown noise matrix parameter is difficult to obtain in nonlinear systems;
(3) For multi-sensor systems with measurement noise signals of the various sensors possibly having different statistics, it is an open issue to estimate and fuse states as well as noise statistics.

Therefore, a novel adaptive filter for a multi-sensor nonlinear system with heavy-tailed measurement noise will be proposed in this paper. Results in this work will help to deal with state estimation and data fusion problems for nonlinear systems with heavy-tailed measurement noise.

3. Variational Bayesian Student's t-Based Cubature Information Filter

3.1. Spherical-Radial Cubature Rule

The main ingredient of the CIF for nonlinear systems is the SRC rule. Based on the prior mean and covariance, an initial set of sampling points are selected and then propagated through the nonlinear function, thus providing transformed sampling points. Then, by weighting these transformed sampling points, the posterior mean and variance can be obtained. Specifically, the SRC rule operates as follows.

Nonlinear filtering involves integrals of this form:

$$\mathrm{Int}(I) = \int_{\mathbb{R}^n} I(x)\ \exp(-x^{\mathrm{T}}x)\,\mathrm{d}x \tag{3}$$

where $I(\cdot)$ is a nonlinear function and $\mathbb{R}^n$ denotes the integration domain. Let $x = rg$ with $g^{\mathrm{T}}g = 1$ and $r \in [0, \infty]$, so that $x^{\mathrm{T}}x = r^2$, where g is a direction vector and r is the radius. Then, integration in (3) can be rewritten as:

$$\mathrm{Int}(I) = \int_0^{\infty} \int_{\Xi} I(rg)\, r^{n-1}\ \exp(-r^2)\,\mathrm{d}\sigma(g)\mathrm{d}r \tag{4}$$

where $\Xi = \{g \in \mathbb{R}^n | g^{\mathrm{T}}g = 1\}$ denotes an n-dimensional unit sphere surface and $\sigma(\cdot)$ is an element on Ξ. Thus, the integration in (4) can be transformed into a spherical integration:

$$\mathrm{Sph}(r) = \int_{\Xi} I(rg)\,\mathrm{d}\sigma(g) \tag{5}$$

and a radial integration:

$$\text{Int} = \int_0^{\infty} \text{Sph}(r) r^{n-1} \exp(-r^2)\, \mathrm{d}r\,. \tag{6}$$

Hence, for a Gaussian integration, it turns out that:

$$\int_{\mathbb{R}^n} I(x)\, \mathcal{N}(x; \mu, \Sigma)\, \mathrm{d}x \cong \frac{1}{2n} \sum_{i=1}^{2n} I(\sqrt{\Sigma} r_i + \mu) \tag{7}$$

where: $\sqrt{\Sigma}$ can be obtained by performing Cholesky decomposition of Σ; $r_i = \sqrt{n}\, \{e\}_i$, $\{e\}_i$ being the i-th column of $\{e\}$, with, for example, $\{e\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$ if $\{e\} \in \mathbb{R}^2$.

3.2. Student's t Distribution and Time Update

The ST distribution [38] can be described as:

$$St(\alpha; 0, M, \kappa) = \int_0^{+\infty} \mathcal{N}(\alpha; 0, M/\gamma)\, G\left(\gamma; \frac{\kappa}{2}, \frac{\kappa}{2}\right) \mathrm{d}\gamma \tag{8}$$

where: $St(\cdot; \mu, M, \kappa)$ represents the ST PDF with mean μ, scale matrix M, and degree of freedom (DOF) parameter κ; γ denotes an auxiliary parameter; $\mathcal{N}(\cdot; \mu, M)$ is the Gaussian PDF with mean μ and variance M; $G(\cdot; a, b)$ denotes the Gamma PDF with parameters a and b.

Since the measurement noise has heavy-tailed characteristic, it can be modeled as an ST distribution with zero mean, scale matrix R_k, and DOF κ_k [38] as parameters, i.e.,

$$St(v_k; 0, R_k, \kappa_k) = \int_0^{+\infty} \mathcal{N}(v_k; 0, R_k/\gamma_k)\, G\left(\gamma_k; \frac{\kappa_k}{2}, \frac{\kappa_k}{2}\right) \mathrm{d}\gamma_k \tag{9}$$

where γ_k denotes the auxiliary parameter at time k.

Prior distributions of R_k, κ_k, as well as γ_k, are selected as inverse Wishart and, respectively, Gamma distribution, i.e.,

$$p(R_{k-1}) = IW(R_{k-1}; \hat{\delta}_{k-1|k-1}, \hat{\Delta}_{k-1|k-1}) \tag{10}$$

$$p(\gamma_{k-1}) = G(\gamma_{k-1}; \hat{a}_{k-1|k-1}, \hat{b}_{k-1|k-1}) \tag{11}$$

$$p(\kappa_{k-1}) = G(\kappa_{k-1}; \hat{\phi}_{k-1|k-1}, \hat{\Phi}_{k-1|k-1}) \tag{12}$$

where $IW(\cdot)$ denotes the inverse Wishart distribution [44] and $\hat{\delta}_{k-1|k-1}$, $\hat{\Delta}_{k-1|k-1}$ are its parameters estimated at time $k-1$. Likewise, $\hat{a}_{k-1|k-1}$, $\hat{b}_{k-1|k-1}$ and $\hat{\phi}_{k-1|k-1}$, $\hat{\Phi}_{k-1|k-1}$ are the corresponding estimated distribution parameters of γ_{k-1} and, respectively, κ_{k-1}.

Time updating is performed by applying a forgetting factor τ to the prior estimates [44], i.e.,

$$\hat{\delta}_{k|k-1} = \tau(\hat{\delta}_{k-1|k-1} - n_z - 1) + n_z + 1 \tag{13}$$

$$\hat{\Delta}_{k|k-1} = \tau \hat{\Delta}_{k-1|k-1} \tag{14}$$

$$\hat{\phi}_{k|k-1} = \tau \hat{\phi}_{k-1|k-1} \tag{15}$$

$$\hat{\Phi}_{k|k-1} = \tau \hat{\Phi}_{k-1|k-1} \tag{16}$$

where: n_z denotes the dimension of the measurement vector y_k; $k|k-1$ stands for prediction from time $k-1$ to k, while $k-1|k-1$ stands for the filtered estimation at $k-1$.

Suppose that the prior state estimate and SECM are $\hat{x}_{k-1|k-1}$ and $P_{k-1|k-1}$, respectively. By applying the SRC rule to the prior information, we have:

$$P_{k-1|k-1} = T_{k-1} T_{k-1}^{\mathrm{T}} \tag{17}$$

$$\chi_{i,k-1} = T_{k-1} r_i + \hat{x}_{k-1|k-1} \quad (i = 1, 2, \ldots, 2n_x) \tag{18}$$

$$\Psi_{i,k-1} = f(\chi_{i,k-1}) \tag{19}$$

where: T_{k-1} is a matrix obtained by performing the Cholesky decomposition of $P_{k-1|k-1}$; r_i represents the i-th column of $\{e\}$ in the SRC rule; $\chi_{i,k-1}$ denotes the i-th sampling point; n_x is the dimension of x_k; $\Psi_{i,k-1}$ denotes the corresponding transformed i-th sampling obtained by propagating $\chi_{i,k-1}$ through $f(\cdot)$. To this end, the predicted $\hat{x}_{k|k-1}$ as well as $P_{k|k-1}$ are obtained as:

$$\hat{x}_{k|k-1} = \frac{1}{2n_x} \sum_{i=1}^{2n_x} \Psi_{i,k-1} \tag{20}$$

$$P_{k|k-1} = \frac{1}{2n_x} \sum_{i=1}^{2n_x} \Psi_{i,k-1} \Psi_{i,k-1}^{\mathrm{T}} - \hat{x}_{k|k-1} \hat{x}_{k|k-1}^{\mathrm{T}} + Q_{k-1} \,. \tag{21}$$

Thus, the predicted information matrix and state vector, $Z_{k|k-1}$ and $\zeta_{k|k-1}$, are obtained through

$$Z_{k|k-1} = P_{k|k-1}^{-1} \tag{22}$$

$$\zeta_{k|k-1} = Z_{k|k-1} \hat{x}_{k|k-1} \,. \tag{23}$$

3.3. Variational Bayesian Student's t-Based Cubature Information Filter (VBST-CIF)

The joint posterior PDF of system state x_k, scale matrix R_k, DOF parameter κ_k, and auxiliary parameter γ_k, denoted as $p(x_k, R_k, \kappa_k, \gamma_k | y_{1:k})$, needs to be computed. However, for the nonlinear system (1) and (2), an analytical solution for $p(x_k, R_k, \kappa_k, \gamma_k | y_{1:k})$ is impossible to obtain. For the purpose of obtaining an approximate solution, the VB approach is introduced.

The key idea of the VB approach is to get an approximate posterior distribution $q(\cdot)$ via minimizing the Kullback–Leibler divergence (KLD) between the approximate $q(\cdot)$ and true one $p(\cdot)$ [36], i.e.,

$$q^* = \arg\min \mathrm{KLD}\,(q||p) \tag{24}$$

where the KLD is defined as:

$$\mathrm{KLD}\,(q||p) \equiv \int q(\cdot) \log \frac{q(\cdot)}{p(\cdot)} \,\mathrm{d}(\cdot) \,. \tag{25}$$

Hence, the problem can be transformed into searching an approximate $q(\cdot)$ that minimizes the KLD between itself and the true $p(\cdot)$, i.e.,

$$q^*(\cdot) = \arg\min \left(q(\cdot) \,||\, p(x_k, R_k, \kappa_k, \gamma_k | y_{1:k}) \right) . \tag{26}$$

The joint posterior distribution of the parameters to be estimated is approximated via the following factored form [36]:

$$p(x_k, R_k, \kappa_k, \gamma_k | y_{1:k}) \cong q_x(x_k)\, q_R(R_k)\, q_\kappa(\kappa_k)\, q_\gamma(\gamma_k) \tag{27}$$

where $q_x(\cdot)$, $q_R(\cdot)$, $q_\kappa(\cdot)$, and $q_\gamma(\cdot)$ represent approximate posterior PDFs for x_k, R_k, κ_k, and γ_k, respectively. By applying logarithm to (27), it can be deduced that:

$$\begin{aligned} \log q(\xi) &= E_{\Omega^{(-\xi)}} (\log p(\Omega, y_{1:k})) + c_\xi \\ \Omega &\equiv \{x_k, R_k, \kappa_k, \gamma_k\} \end{aligned} \tag{28}$$

where: Ω is the set of parameters $(x_k, R_k, \kappa_k, \gamma_k)$ that need to be estimated; ξ is an element of Ω; $\Omega^{(-\xi)}$ denotes the complementary set of ξ in Ω; $E(\cdot)$ denotes the expectation operator; c_ξ is a quantity that is independent of ξ.

By exploiting the hierarchical Gaussian form in [38], it can be obtained that:

$$\begin{aligned} &p(x_k, R_k, \kappa_k, \gamma_k, y_{1:k}) \\ &= \mathcal{N}\left(y_k; h(x_k), \frac{R_k}{\gamma_k}\right) \mathcal{N}\left(x_k; \hat{x}_{k|k-1}, P_{k|k-1}\right) IW(R_k; \delta_k, \Delta_k)\, G\left(\gamma_k; \frac{\kappa_k}{2}, \frac{\kappa_k}{2}\right) G\left(\kappa_k; \phi_k, \Phi_k\right) p(y_{1:k-1}) \end{aligned} \tag{29}$$

Setting $\xi = R_k$,

$$\begin{aligned} \log q^{(j+1)}(\xi) &= \log q^{(j+1)}(R_k) \\ &= -0.5\,(\hat{\delta}_{k|k-1} + n_z + 2)\log(R_k) - 0.5\,\mathrm{tr}\left\{[\hat{\Delta}_{k|k-1} + E^{(j+1)}(\gamma_k)\, B_k^{(j)}]\, R_k^{-1}\right\} + c_R \end{aligned} \tag{30}$$

from which we have:

$$q^{(j+1)}(R_k) = IW\left(R_k; \hat{\delta}_{k|k}^{(j+1)}, \hat{\Delta}_{k|k}^{(j+1)}\right) \tag{31}$$

whose parameters $\hat{\delta}_{k|k}^{(j+1)}$ and $\hat{\Delta}_{k|k}^{(j+1)}$ can be obtained by:

$$\hat{\delta}_{k|k}^{(j+1)} = \hat{\delta}_{k|k-1} + 1 \tag{32}$$

$$\hat{\Delta}_{k|k}^{(j+1)} = \hat{\Delta}_{k|k-1} + E^{(j+1)}(\gamma_k) B_k^{(j)} \tag{33}$$

where $B_k^{(j)}$ can be obtained via the SRC rule as follows:

$$P_{k|k}^{(j)} = T_k T_k^{\mathrm{T}} \tag{34}$$

$$\chi_{i,k}^{(j)} = T_k r_i + \hat{x}_{k|k}^{(j)} \tag{35}$$

$$\Upsilon_{i,k}^{(j)} = h(\chi_{i,k}^{(j)}) \tag{36}$$

$$\hat{y}_{k|k}^{(j)} = \frac{1}{2n_x} \sum_{i=1}^{2n_x} \Upsilon_{i,k}^{(j)} \tag{37}$$

$$B_k^{(j)} = \frac{1}{2n_x} \sum_{i=1}^{2n_x} \Upsilon_{i,k}^{(j)} \Upsilon_{i,k}^{(j)\mathrm{T}} - \hat{y}_{k|k}^{(j)} (\hat{y}_{k|k}^{(j)})^{\mathrm{T}}. \tag{38}$$

Next, setting $\xi = \gamma_k$, we get:

$$\begin{aligned} \log q^{(j+1)}(\xi) &= \log q^{(j+1)}(\gamma_k) \\ &= \left(\frac{n_z + E^{(j)}(\kappa_k)}{2} - 1\right) \log(\gamma_k) - 0.5\left\{\mathrm{tr}\left[B_k^{(j)} E^{(j)}(R_k^{-1})\right] + E^{(j)}(\kappa_k)\right\}\gamma_k + c_\gamma \end{aligned} \tag{39}$$

from which:

$$q^{(j+1)}(\gamma_k) = G\left(\gamma_k; \hat{a}_{k|k}^{(j+1)}, \hat{b}_{k|k}^{(j+1)}\right) \tag{40}$$

where:

$$\hat{a}_{k|k}^{(j+1)} = 0.5\,(n_z + E^{(j)}(\kappa_k)) \tag{41}$$

$$\hat{b}_{k|k}^{(j+1)} = -0.5\left\{\mathrm{tr}\left[B_k^{(j)} E^{(j)}(R_k^{-1})\right] + E^{(j)}(\kappa_k)\right\}. \tag{42}$$

Then, setting $\xi = \kappa_k$, we get

$$\begin{aligned}\log q^{(j+1)}(\xi) =& \log q^{(j+1)}(\kappa_k) \\ =& \frac{\kappa_k}{2}\log\left(\frac{\kappa_k}{2}\right) - \log\Gamma\left(\frac{\kappa_k}{2}\right) \\ &+ \left(\frac{\kappa_k}{2}-1\right) E^{(j+1)}(\log(\gamma_k)) - \frac{\kappa_k}{2}E^{(j+1)}(\gamma_k) + (\hat{\phi}_{k|k-1}-1)\log(\kappa_k) - \hat{\Phi}_{k|k-1}\kappa_k + c_\kappa \end{aligned} \tag{43}$$

from which we have:

$$q^{(j+1)}(\kappa_k) = G\left(\kappa_k; \hat{\phi}_{k|k}^{(j+1)}, \hat{\Phi}_{k|k}^{(j+1)}\right) \tag{44}$$

where:

$$\hat{\phi}_{k|k}^{(j+1)} = \hat{\phi}_{k|k-1} + 0.5 \tag{45}$$

$$\hat{\Phi}_{k|k}^{(j+1)} = \hat{\Phi}_{k|k-1} - 0.5E^{(j+1)}(\log(\gamma_k)) + 0.5E^{(j+1)}(\gamma_k) - 0.5. \tag{46}$$

Finally, setting $\xi = x_k$,

$$\begin{aligned}\log q^{(j+1)}(\xi) =& \log q^{(j+1)}(x_k) \\ =& -0.5(x_k - \hat{x}_{k|k-1})^{\mathrm{T}} E^{(j)}(P_{k|k-1}^{-1})(x_k - \hat{x}_{k|k-1}) \\ &- 0.5\, E^{(j)}(\gamma_k)(y_k - h(x_k))^{\mathrm{T}} E^{(j)}(R_k^{-1})(y_k - h(x_k)) + c_x \end{aligned} \tag{47}$$

according to which we have:

$$q^{(j+1)}(x_k) = \mathcal{N}\left(x_k; \hat{x}_{k|k}^{(j+1)}, P_{k|k}^{(j+1)}\right) \tag{48}$$

where $\hat{x}_{k|k}^{(j+1)}$ and $P_{k|k}^{(j+1)}$ are the estimated state and SECM in the j-th VB iteration, respectively.

In order to obtain $\hat{x}_{k|k}^{(j+1)}$ and $P_{k|k}^{(j+1)}$, the associated information matrix and information vector contributions $I_k^{(j+1)}$ and, respectively, $\iota_k^{(j+1)}$ should be first computed. By utilizing a linearized error propagation method [26], they can be obtained as follows:

$$I_k^{(j+1)} = H_k^{\mathrm{T}} (\tilde{R}_{k|k}^{(j+1)})^{-1} H_k \tag{49}$$

$$\iota_k^{(j+1)} = H_k^{\mathrm{T}} (\tilde{R}_{k|k}^{(j+1)})^{-1} (y_k - \hat{y}_{k|k-1} + H_k \hat{x}_{k|k-1}) \tag{50}$$

where $\tilde{R}_{k|k}^{(j+1)}$ denotes the estimated scale matrix and H_k a pseudo-measurement matrix. Then, exploiting (31) and a property of the inverse Wishart distribution, it can be obtained that:

$$E^{(j+1)}(R_k^{-1}) = (\hat{\delta}_{k|k}^{(j+1)} - n_z - 1)(\hat{\Delta}_{k|k}^{(j+1)})^{-1}. \tag{51}$$

Similarly, taking into account (40) and (44) as well as a property of the Gamma distribution, we have:

$$E^{(j+1)}(\kappa_k) = \hat{\phi}_{k|k}^{(j+1)} / \hat{\Phi}_{k|k}^{(j+1)} \tag{52}$$

$$E^{(j+1)}(\gamma_k) = \hat{a}_{k|k}^{(j+1)} / \hat{b}_{k|k}^{(j+1)}. \tag{53}$$

Then, $\tilde{R}_{k|k}^{(j+1)}$ is computed as:

$$\tilde{R}_{k|k}^{(j+1)} = (\hat{\delta}_{k|k}^{(j+1)} - n_z - 1)^{-1} \hat{\Delta}_{k|k}^{(j+1)}. \tag{54}$$

Moreover, the pseudo-measurement matrix H_k in (49) and (50) can be obtained as:

$$H_k = P_{xz,k|k-1}^{\mathrm{T}} Z_{k|k-1} \tag{55}$$

where $P_{xz,k|k-1}$ is computed by means of the SRC rule as follows:

$$P_{k|k-1} = T_{k|k-1} T_{k|k-1}^{\mathrm{T}} \tag{56}$$

$$\chi_{i,k|k-1} = T_{k|k-1} r_i + \hat{x}_{k|k-1} \tag{57}$$

$$\Upsilon_{i,k|k-1} = h(\chi_{i,k|k-1}) \tag{58}$$

$$\hat{y}_{k|k-1} = \frac{1}{2n_x} \sum_{i=1}^{2n_x} \Upsilon_{i,k|k-1} \tag{59}$$

$$P_{xz,k|k-1} = \frac{1}{2n_x} \sum_{i=1}^{2n_x} \chi_{i,k-1} \Upsilon_{i,k-1}^{\mathrm{T}} - \hat{x}_{k|k-1} \hat{y}_{k|k-1}^{\mathrm{T}} . \tag{60}$$

Then, the information matrix and vector are updated as follows:

$$Z_{k|k}^{(j+1)} = Z_{k|k-1} + I_k^{(j+1)} \tag{61}$$

$$\zeta_{k|k}^{(j+1)} = \zeta_{k|k-1} + \iota_k^{(j+1)} . \tag{62}$$

Finally, we have:

$$\hat{x}_{k|k}^{(j+1)} = Z_{k|k}^{(j+1)} \backslash \zeta_{k|k}^{(j+1)} \tag{63}$$

$$P_{k|k}^{(j+1)} = Z_{k|k}^{(j+1)} \backslash I_{n_x} \tag{64}$$

where I_{n_x} denotes the $n_x \times n_x$ identity matrix. To summarize the above developments, the pseudocode of VBST-CIF is reported in Algorithm 1.

Algorithm 1 Time-recursion of VBST-CIF

Input: $\hat{x}_{k-1|k-1}$, $P_{k-1|k-1}$, $\hat{\delta}_{k-1|k-1}$, $\hat{\Delta}_{k-1|k-1}$, $\hat{\phi}_{k-1|k-1}$, $\hat{\Phi}_{k-1|k-1}$, y_k, Q_{k-1}, τ

Step 1: Time update:

(1) Compute $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ based on the SRC rule by (17)–(21).

(2) Compute $Z_{k|k-1}$ and $\zeta_{k|k-1}$ by (22) and (23).

(3) Compute $\hat{\delta}_{k|k-1}$ and $\hat{\Delta}_{k|k-1}$ by (13) and (14).

(4) Compute $\hat{\phi}_{k|k-1}$ and $\hat{\Phi}_{k|k-1}$ by (15) and (16).

Step 2: Variational fixed point iterations:

Initialize:

$\hat{x}_{k|k}^{(0)} = \hat{x}_{k|k-1}$, $P_{k|k}^{(0)} = P_{k|k-1}$, $\hat{\delta}_{k|k}^{(0)} = \hat{\delta}_{k|k-1}$, $\hat{\Delta}_{k|k}^{(0)} = \hat{\Delta}_{k|k-1}$, $\hat{\phi}_{k|k}^{(0)} = \hat{\phi}_{k|k-1}$, $\hat{\Phi}_{k|k}^{(0)} = \hat{\Phi}_{k|k-1}$

Measurement update:

(1) Perform Cholesky decomposition of $P_{k|k-1}$ and cubature sampling via (56) and (57) .

(2) Compute $P_{xz,k|k-1}$ by (58)–(60).

(3) Variational parameter update:

for $j = 0, 1, \ldots, N-1$ **do**

(a) Compute $q^{(j+1)}(R_k)$ by (31):

Compute $B_k^{(j)}$ by (34)–(38).

Update $\hat{\delta}_{k|k}^{(j+1)}$ and $\hat{\Delta}_{k|k}^{(j+1)}$ by (32) and (33).

(b) Compute $q^{(j+1)}(\gamma_k)$ by (40):

Compute $\hat{a}_{k|k}^{(j+1)}$ and $\hat{b}_{k|k}^{(j+1)}$ by (41) and (42).

(c) Compute $q^{(j+1)}(\kappa_k)$ by (44):

Update $\hat{\phi}_{k|k}^{(j+1)}$ and $\hat{\Phi}_{k|k}^{(j+1)}$ by (45) and (46).

(d) Compute $q^{(j+1)}(x_k)$ by (48):

Compute $\tilde{R}_{k|k}^{(j+1)}$ by (54).

Compute H_k by (55).

Update $Z_{k|k}^{(j+1)}$ and $\zeta_{k|k}^{(j+1)}$ by (61) and (62).

Compute $\hat{x}_{k|k}^{(j+1)}$ and $P_{k|k}^{(j+1)}$ by (63) and (64).

end for

Step 3: State update:

$$\hat{x}_{k|k} = \hat{x}_{k|k}^{(N)}, P_{k|k} = P_{k|k}^{(N)}, \hat{\delta}_{k|k} = \hat{\delta}_{k|k}^{(N)}, \hat{\Delta}_{k|k} = \hat{\Delta}_{k|k}^{(N)}, \hat{\phi}_{k|k} = \hat{\phi}_{k|k}^{(N)}, \hat{\Phi}_{k|k} = \hat{\Phi}_{k|k}^{(N)}$$

Output: $\hat{x}_{k|k}, P_{k|k}, \hat{\delta}_{k|k}, \hat{\Delta}_{k|k}, \hat{\phi}_{k|k}, \hat{\Phi}_{k|k}$

4. Variational Bayesian Student's t-Based Cubature Information Feedback Fusion (VBST-CIFF)

4.1. Multi-Sensor Cubature Information Feedback Fusion (CIFF)

The information feedback fusion (IFF) algorithm derived in this section is suitable for multi-sensor fusion. The overall structure of the IFF is schematized in Figure 1, where each local filter receives raw measurements from the relative sensor and produces sensor-dependent variables to be used by the information fusion center in order to update the information matrix and state vector, which are then fed back to the local filters.

Based on the structure of IFF, the CIFF algorithm is derived as follows. Consider that there are S sensors. For each sensor s, $s \in \{1, 2, \ldots, S\}$, a measurement $y_{k,s}$ is available at time k, i.e.,

$$y_{k,s} = h_s(x_k) + v_{k,s} \tag{65}$$

where $h_s(\cdot)$ denotes the measurement function of sensor s, and $v_{k,s}$ the corresponding measurement noise. It is assumed that $v_{k,m}$ and $v_{k,n}$ are uncorrelated for any $m \neq n$. Let, at time $k-1$, the state estimate and SECM for all sensors be $\hat{x}_{k-1|k-1}$ and, respectively, $P_{k-1|k-1}$. Correspondingly, the initial information vector and matrix for all sensors are $\zeta_{k-1|k-1}$ and $Z_{k-1|k-1}$, respectively. Then, by exploiting the SRC rule, predicted $\zeta_{k-1|k}$ and $Z_{k-1|k}$ can be obtained, i.e.,

$$(Z_{k-1|k-1})^{-1} = T_{k-1} T_{k-1}^{\mathrm{T}} \tag{66}$$

where $Z_{k-1|k-1}$ is the inverse of $P_{k-1|k-1}$. Transforming $\chi_{i,k-1}$ in (18) through $f(\cdot)$, we have $\Psi_{i,k-1}$ in (19). Then, predicted $Z_{k|k-1}$ and $\zeta_{k|k-1}$, are computed via (22) and (23).

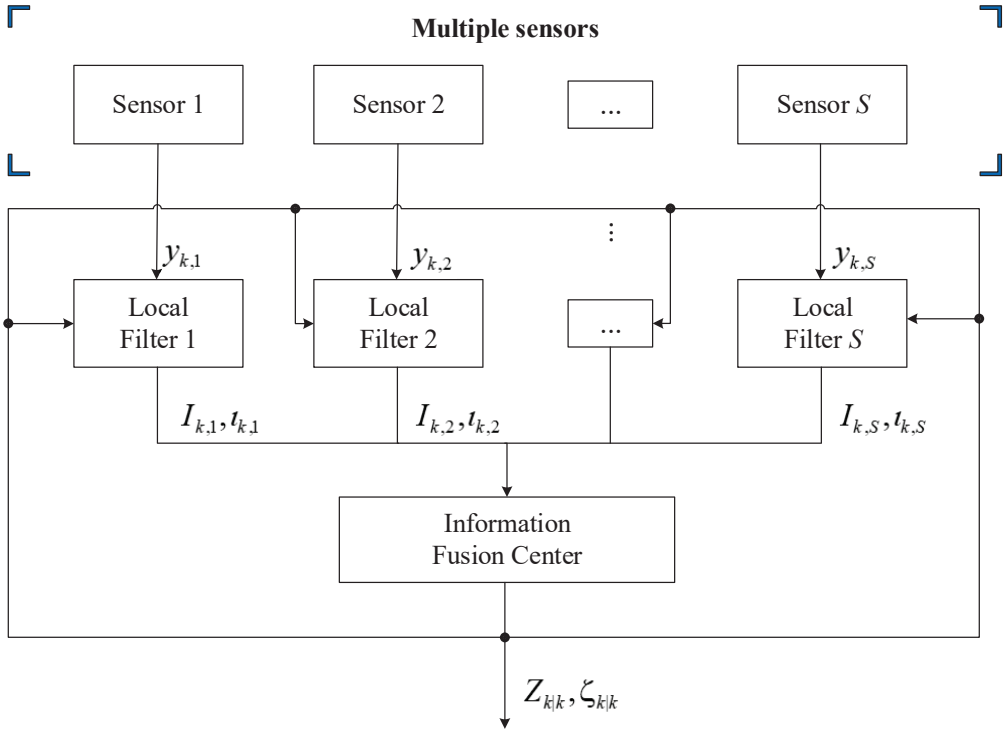

Figure 1. Multi-sensor information feedback fusion.

At the fusion center, the estimated information vector and matrix can, therefore, be updated as follows:

$$\zeta_{k|k} = \zeta_{k|k-1} + \sum_{s=1}^{S} \iota_{k,s} \tag{67}$$

$$Z_{k|k} = Z_{k|k-1} + \sum_{s=1}^{S} I_{k,s} \tag{68}$$

where: the local sensor contributions $I_{k,s}$ and $\iota_{k,s}$ are given by:

$$I_{k,s} = H_{k,s}^{\mathrm{T}} R_{k,s}^{-1} H_k \tag{69}$$

$$\iota_{k,s} = H_{k,s}^{\mathrm{T}} R_{k,s}^{-1} (y_{k,s} - \hat{y}_{k|k-1,s} + H_k \hat{x}_{k|k-1})\,; \tag{70}$$

the pseudo-measurement matrix $H_{k,s}$ of sensor s can be obtained through the statistical linearized error propagation approach as in Section 3.3; $R_{k,s}$ denotes the MNCM of sensor s; $\hat{y}_{k|k-1,s}$ can be obtained by the SRC rule as in the conventional CIF [25].

4.2. Variational Bayesian Student's t-Based Cubature Information Feedback Fusion (VBST-CIFF)

Consider now that each sensor s is characterized by heavy-tailed measurement noise $v_{k,s}$. In the time-update, the parameters $\hat{\delta}_{k|k-1,s}$, $\hat{\Delta}_{k|k-1,s}$, $\hat{\phi}_{k|k-1,s}$, and $\hat{\Phi}_{k|k-1,s}$ for sensor s, can be obtained via (71)–(74), respectively, i.e.,

$$\hat{\delta}_{k|k-1,s} = \tau(\hat{\delta}_{k-1|k-1,s} - n_z - 1) + n_z + 1 \tag{71}$$

$$\hat{\Delta}_{k|k-1,s} = \tau \hat{\Delta}_{k-1|k-1,s} \tag{72}$$

$$\hat{\phi}_{k|k-1,s} = \tau \hat{\phi}_{k-1|k-1,s} \tag{73}$$

$$\hat{\Phi}_{k|k-1,s} = \tau \hat{\Phi}_{k-1|k-1,s}\,. \tag{74}$$

Then, in the measurement update, the estimated scale matrix for sensor s is given by:

$$\tilde{R}_{k|k,s}^{(j+1)} = \left(\hat{\delta}_{k|k,s}^{(j+1)} - n_z - 1\right)^{-1} \hat{\Delta}_{k|k,s}^{(j+1)} \tag{75}$$

where $\hat{\delta}_{k|k,s}^{(j+1)}$ and $\hat{\Delta}_{k|k,s}^{(j+1)}$ can be obtained via:

$$\hat{\delta}_{k|k,s}^{(j+1)} = \hat{\delta}_{k|k-1} + 1 \tag{76}$$

$$\hat{\Delta}_{k|k,s}^{(j+1)} = \hat{\Delta}_{k|k-1} + E^{(j+1)}(\gamma_{k,s}) B_{k,s}^{(j)} \tag{77}$$

where $B_{k,s}^{(j)}$ can be derived in the same way as in (34)–(38) with cubature sampling points transformed through $h_s(\cdot)$ as:

$$\Upsilon_{i,k|k-1,s} = h_s(\chi_{i,k|k-1,s}) \tag{78}$$

$$\hat{y}_{k|k-1,s} = \frac{1}{2n_z} \sum_{i=1}^{2n_z} \Upsilon_{i,k|k-1,s} \tag{79}$$

in which the original sampling points $\chi_{i,k|k-1,s}$ are given by (56) and (57). Furthermore, $E^{(j+1)}(\gamma_{k,s})$ for sensor s can be obtained by:

$$E^{(j+1)}(\gamma_{k,s}) = \hat{a}_{k|k,s}^{(j+1)} / \hat{b}_{k|k,s}^{(j+1)} \tag{80}$$

where,

$$\hat{a}_{k|k,s}^{(j+1)} = 0.5\left(n_z + E^{(j)}(\kappa_{k,s})\right) \tag{81}$$

$$\hat{b}_{k|k,s}^{(j+1)} = -0.5\left\{\mathrm{tr}\left[B_{k,s}^{(j+1)} E^{(j)}(R_{k,s}^{-1})\right] + E^{(j)}(\kappa_{k,s})\right\} \tag{82}$$

and $E^{(j+1)}(\kappa_{k,s})$ is given by:

$$E^{(j+1)}(\kappa_{k,s}) = \hat{\phi}_{k|k,s}^{(j+1)} / \hat{\Phi}_{k|k,s}^{(j+1)} \tag{83}$$

in which:

$$\hat{\phi}_{k|k,s}^{(j+1)} = \hat{\phi}_{k|k-1,s} + 0.5 \tag{84}$$

$$\hat{\Phi}_{k|k,s}^{(j+1)} = \hat{\Phi}_{k|k-1,s} - 0.5E^{(j+1)}(\log(\gamma_{k,s})) + 0.5E^{(j+1)}(\gamma_{k,s}) - 0.5\,. \tag{85}$$

The pseudo-measurement matrix for sensor s turns out to be:

$$H_{k,s} = P_{xz,k|k-1,s}^{\mathrm{T}} Z_{k|k-1} \tag{86}$$

where $P_{xz,k|k-1,s}^{\mathrm{T}}$ can be obtained via

$$P_{xz,k|k-1,s} = \frac{1}{2n_x} \sum_{i=1}^{2n_x} \chi_{i,k-1} \Upsilon_{i,k-1,s}^{\mathrm{T}} - \hat{x}_{k|k-1} \hat{y}_{k|k-1,s}^{\mathrm{T}}\,. \tag{87}$$

Hence, the information matrix and vector contributions due to sensor s can be derived as follows:

$$I_{k,s}^{(j+1)} = H_{k,s}^{\mathrm{T}} \left(\tilde{R}_{k|k,s}^{(j+1)} \right)^{-1} H_{k,s} \tag{88}$$

$$\iota_{k,s}^{(j+1)} = H_{k,s}^{\mathrm{T}} \left(\tilde{R}_{k|k,s}^{(j+1)} \right)^{-1} \left(y_{k,s} - \hat{y}_{k|k-1,s} + H_{k,s} \hat{x}_{k|k-1} \right) \tag{89}$$

from which the information matrix and vector can be updated according to

$$Z_{k|k}^{(j+1)} = Z_{k|k-1} + \sum_{s=1}^{S} I_{k,s}^{(j+1)} \tag{90}$$

$$\zeta_{k|k}^{(j+1)} = \zeta_{k|k-1} + \sum_{s=1}^{S} \iota_{k,s}^{(j+1)} . \tag{91}$$

Finally, the global state estimate and covariance can be obtained by (63) and (64). To summarize the above developments, the flow-chart of the proposed VBST-CIFF is shown in Figure 2. Accordingly, the pseudocode is reported in Algorithm 2.

Please notice that VBST-CIFF is actually tightly-coupled (centralized) in that the raw measurements from all sensors are integrated into a single nonlinear information filter. Thanks to the adoption of the information filter form and the assumption that measurement noise signals of the various sensors are uncorrelated, the filter's multi-sensor measurement update turns out to be decoupled into single-sensor updates as in (67) and (68). However, the overall filter is coupled through the information feedback of Z_k, ζ_k into the local filters (see Figure 1).

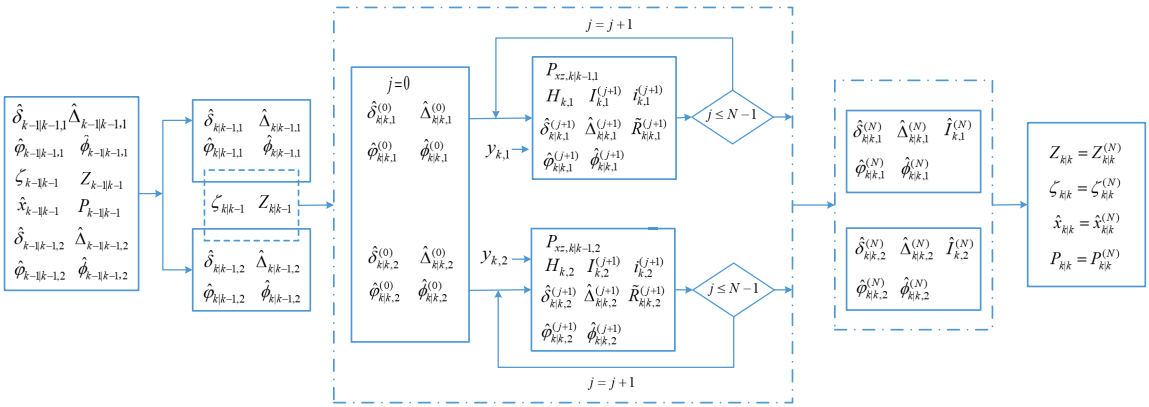

Figure 2. Flow-chart of proposed VBST-CIFF (variational Bayesian Student's t-based-cubature information feedback fusion).

Algorithm 2 Time-recursion of VBST-CIFF

Input: $\hat{x}_{k-1|k-1}$, $P_{k-1|k-1}$, $y_{k,s}$, τ, $\hat{\delta}_{k-1|k-1,s}$, $\hat{\Delta}_{k-1|k-1,s}$, $\hat{\phi}_{k-1|k-1,s}$, $\hat{\Phi}_{k-1|k-1,s}$, Q_{k-1}

Step 1: Time update:

(1) Compute $Z_{k|k-1}$ and $\zeta_{k|k-1}$ by (22) and (23).

(2) Compute $\hat{\delta}_{k|k-1,s}$ and $\hat{\Delta}_{k|k-1,s}$ by (71) and (72).

(3) Compute $\hat{\phi}_{k|k-1,s}$ and $\hat{\Phi}_{k|k-1,s}$ by (73) and (74).

Step 2: Variational fixed-point iterations:

Initialize:

$$\hat{x}^{(0)}_{k|k} = \hat{x}_{k|k-1},\ P^{(0)}_{k|k} = P_{k|k-1},\ \hat{\delta}^{(0)}_{k|k,s} = \hat{\delta}_{k|k-1,s},\ \hat{\Delta}^{(0)}_{k|k,s} = \hat{\Delta}_{k|k-1,s},\ \hat{\phi}^{(0)}_{k|k,s} = \hat{\phi}_{k|k-1,s},\ \hat{\Phi}^{(0)}_{k|k,s} = \hat{\Phi}_{k|k-1,s}$$

Measurement update:

(1) Perform Cholesky decomposition of $P_{k|k-1}$ and cubature sampling via (56) and (57) .

(2) Compute $P_{xz,k|k-1,s}$ by (87).

(3) Variational parameter update:

for $j = 0, 1, \ldots, N-1$ **do**

(a) Update $E^{(j)}(\kappa_{k,s})$, $\hat{\phi}^{(j+1)}_{k|k,s}$, and $\hat{\Phi}^{(j+1)}_{k|k,s}$ by (83)–(85).

(b) Update $E^{(j+1)}(\gamma_{k,s})$, $\hat{a}^{(j+1)}_{k|k,s}$, and $\hat{b}^{(j+1)}_{k|k,s}$ by (80)–(82).

(c) Update $\tilde{R}^{(j+1)}_{k|k,s}$, $\hat{\delta}^{(j+1)}_{k|k,s}$ and $\hat{\Delta}^{(j+1)}_{k|k,s}$ by (75)–(77).

(d) Update $H_{k,s}$ by (86).

(e) Update $I^{(j+1)}_{k,s}$ and $\iota^{(j+1)}_{k,s}$ by (88) and (89).

(f) Set variational parameters:

$$\hat{\delta}_{k|k,s} = \hat{\delta}^{(N)}_{k|k,s},\ \hat{\Delta}_{k|k,s} = \hat{\Delta}^{(N)}_{k|k,s},\ \hat{\phi}_{k|k,s} = \hat{\phi}^{(N)}_{k|k,s},\ \hat{\Phi}_{k|k,s} = \hat{\Phi}^{(N)}_{k|k,s}$$

end for

Step 3: Fusion (at the global center):

(1) Update $Z^{(N)}_{k|k}$ and $\zeta^{(N)}_{k|k}$ by (90) and (91);

(2) Update $\hat{x}^{(N)}_{k|k}$ and $P^{(N)}_{k|k}$ via (63) and (64).

(3) Set:

$$\hat{x}_{k|k} = \hat{x}^{(N)}_{k|k},\ P_{k|k} = P^{(N)}_{k|k},\ \hat{\delta}_{k|k,s} = \hat{\delta}^{(N)}_{k|k,s},\ \hat{\Delta}_{k|k,s} = \hat{\Delta}^{(N)}_{k|k,s},\ \hat{\phi}_{k|k} = \hat{\phi}^{(N)}_{k|k,s},\ \hat{\Phi}_{k|k} = \hat{\Phi}^{(N)}_{k|k,s}$$

Output: $\hat{x}_{k|k}$, $P_{k|k}$, $\hat{\delta}_{k|k,s}$, $\hat{\Delta}_{k|k,s}$, $\hat{\phi}_{k|k,s}$, $\hat{\Phi}_{k|k,s}$

5. Simulation Results

5.1. VBST-CIF Single-Sensor Target Tracking

Consider a target car tracking scenario with just a single sensor (a 2D radar). The state of the car is taken as $x = [\varsigma, \dot{\varsigma}, \mu, \dot{\mu}, \omega]^{\mathrm{T}}$, where ς and μ are the Cartesian coordinates of position, while $\dot{\varsigma}$ and $\dot{\mu}$ represent the corresponding velocity, and ω is the turning rate. The car moves according to the coordinated turn (CT) model:

$$x_k = F(\omega_{k-1})x_{k-1} + w_{k-1} \tag{92}$$

with a turning-rate dependent state-transition matrix:

$$F(\omega) = \begin{bmatrix} 1 & \frac{\sin(\omega)}{\omega} & 0 & -\frac{1-\cos(\omega)}{\omega} & 0 \\ 0 & \cos(\omega) & 0 & -\sin(\omega) & 0 \\ 0 & \frac{1-\cos(\omega)}{\omega} & 1 & \frac{\sin(\omega)}{\omega} & 0 \\ 0 & \sin(\omega) & 0 & \cos(\omega) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

fixed at turning rate $\omega = 10°$; $w_{k-1} = \mathcal{N}(0, Q_{k-1})$;

$$Q_{k-1} = \begin{bmatrix} \frac{p_1 T^3}{3} & \frac{p_1 T^2}{2} & 0 & 0 & 0 \\ \frac{p_1 T^2}{2} & T & 0 & 0 & 0 \\ 0 & 0 & \frac{p_1 T^3}{3} & \frac{p_1 T^2}{2} & 0 \\ 0 & 0 & \frac{p_1 T^2}{2} & T & 0 \\ 0 & 0 & 0 & 0 & p_2 T \end{bmatrix},$$

with noise intensity factors $p_1 = 0.1$, $p_2 = 1.75 \times 10^{-4}$ and sampling interval $T = 1\,[s]$. Conversely, the measurement function of the radar is:

$$y_k = \begin{bmatrix} \sqrt{\varsigma_k^2 + \mu_k^2} \\ \text{atan2}\left(\varsigma_k, \mu_k\right) \end{bmatrix} + v_k \tag{93}$$

where: $\text{atan2}(\varsigma, \mu)$ denotes the 4-quadrant inverse-tangent defined as the argument of the complex number $\varsigma + j\mu$, j denoting the imaginary unit; v_k is heavy-tailed measurement noise such that:

$$v_k \sim \begin{cases} \mathcal{N}(0, \Lambda_k), & \text{with probability } 0.9 \\ \mathcal{N}(0, 100\Lambda_k), & \text{with probability } 0.1 \end{cases} \tag{94}$$

with $\Lambda_k = \begin{bmatrix} 4\,[\text{m}^2] & 0 \\ 0 & 10^{-4}\,[\text{rad}^2] \end{bmatrix}$. Moreover, the initial values for state and SECM of the car are set as follows:

$$x_0 = [100\,[\text{m}], 3\,[\text{m/s}], 100\,[\text{m}], 2\,[\text{m/s}], \frac{-10\pi}{180}\,[\text{rad}]\,]^{\text{T}} \tag{95}$$

$$P_0 = \begin{bmatrix} 10\,[\text{m}^2] & 0 & 0 & 0 & 0 \\ 0 & 1\,[\text{m}^2/s^2] & 0 & 0 & 0 \\ 0 & 0 & 10\,[\text{m}^2] & 0 & 0 \\ 0 & 0 & 0 & 1\,[\text{m}^2/s^2] & 0 \\ 0 & 0 & 0 & 0 & 10^{-4}\,[\text{rad}^2] \end{bmatrix}. \tag{96}$$

A total of 100 independent Monte Carlo trials have been carried out, with each trial lasting for $t = 50\,[\text{s}]$. Furthermore, the root-mean-square-error (RMSE) is adopted to assess tracking precision. For variable b at time k, it is defined as:

$$\text{RMSE}_b(k) = \sqrt{\frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} \left(b_k^i - \hat{b}_k^i\right)^{\text{T}} \left(b_k^i - \hat{b}_k^i\right)} \tag{97}$$

where: N_{MC} denotes the number of Monte Carlo trials; b denotes either the position $p = [\varsigma, \mu]^{\text{T}}$ or velocity $v = [\dot{\varsigma}, \dot{\mu}]^{\text{T}}$; and b_k^i is the true value in the i-th trial at time k. Conversely, $\hat{b}_k^i$ represents the estimated value in the i-th trial at time k. Furthermore, the mean-root-mean-square-error (MRMSE)

is used to measure the average RMSE during the whole simulation time. Specifically, MRMSE for variable b is defined as:

$$\mathrm{RMSE}_b = \frac{1}{t/T} \sum_{k=1}^{t/T} \mathrm{RMSE}_b(k). \tag{98}$$

Simulation results are presented to compare the conventional CIF [43] with the proposed VBST-CIF. Simulation parameters have been chosen in the same way as in [33]. Specifically, in CIF, the process noise covariance Q and MNCM R are fixed to $Q = Q_{k-1}$ and $R = \Lambda_k$. Conversely, in the proposed VBST-CIF, $Q = Q_{k-1}$, while the scale matrix as well as other noise statistical parameters are adaptively estimated. Other parameters are set in the same way as in [33]: forgetting the factor $\tau = 0.9$; number of VB iterations $N = 10$; initial noise parameters of the scale matrix $\delta = n_z + 2 = 4$ and $\Delta = \Lambda_k$; initial noise parameters of the DOF parameter $\phi = 5$ and $\Phi = 1$.

Results are presented in Figures 3–6 and Table 1. Figures 3–6 demonstrate that the proposed VBST-CIF outperforms conventional CIF in position, velocity, and turning-rate estimation. Precisely, Table 1 shows that VBST-CIF yields an MRMSE improvement with respect to conventional CIF of respectively 56% in position, 48% in velocity, and 20% in turning-rate.

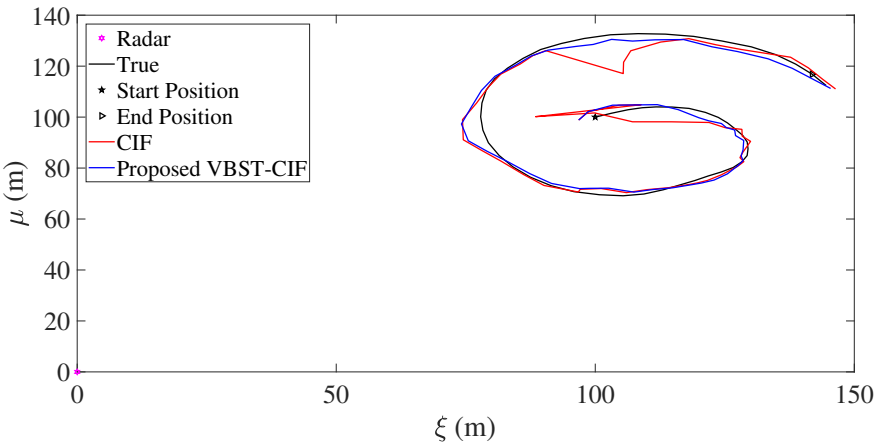

Figure 3. Target trajectory.

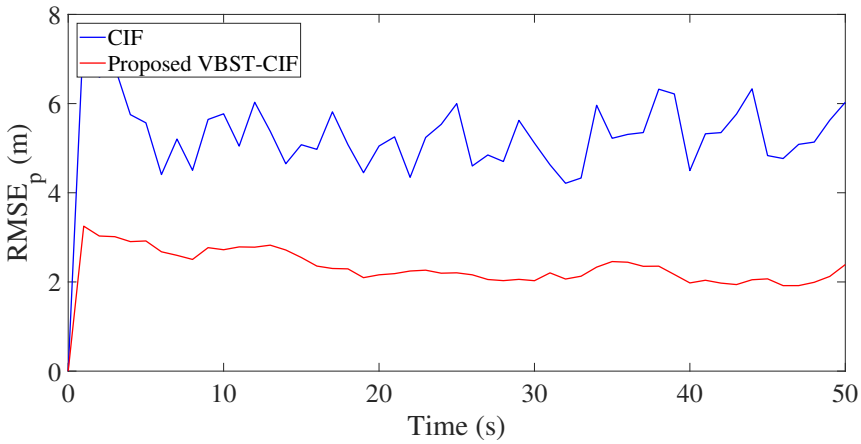

Figure 4. Position root-mean-square-error (RMSE) in single-sensor target tracking.

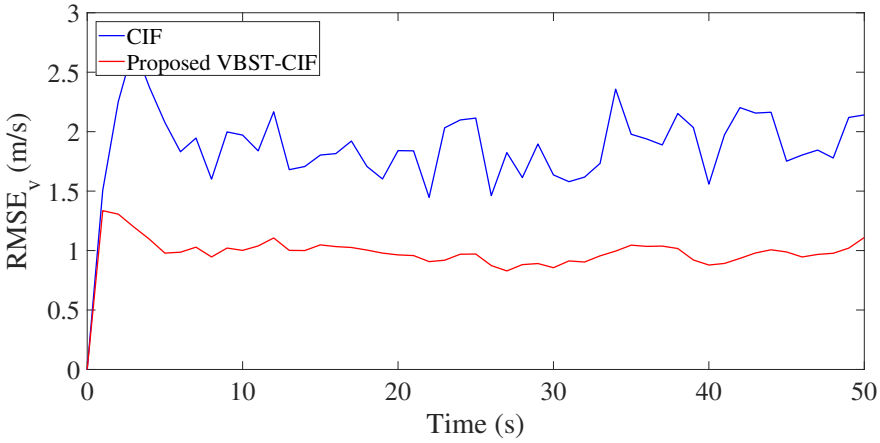

Figure 5. Velocity RMSE in single-sensor target tracking.

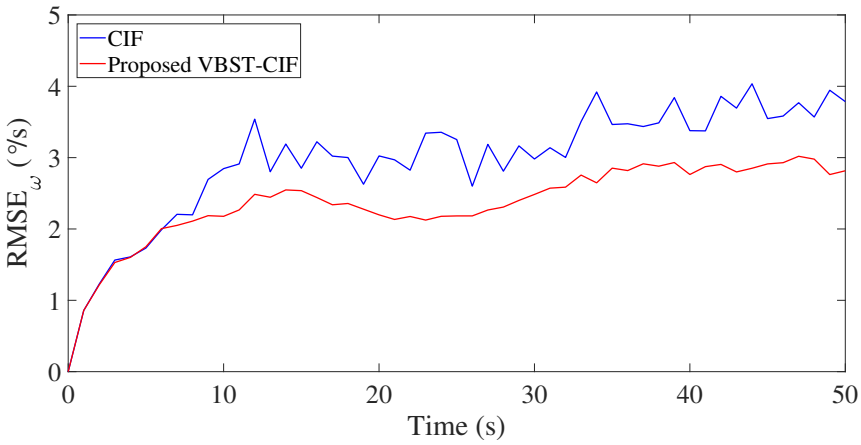

Figure 6. Turning-rate RMSE in single-sensor target tracking.

Table 1. Mean-root-mean-square-error (MRMSE) in single-sensor target tracking.

Filter	$\mathbf{RMSE}_p$	$\mathbf{RMSE}_v$	$\mathbf{RMSE}_\omega$
CIF	5.2364	1.8646	0.0518
VBST-CIF	2.3052	0.9740	0.0412

5.2. VBST-CIFF Multi-Sensor Target Tracking

Let us now consider a multi-sensor target (car) tracking scenario with two (radar) sensors S_1 and S_2, located at $(10, 10)$ and respectively at the origin $(0, 0)$. The car's state is taken as $x = [\varsigma, \dot{\varsigma}, \mu, \dot{\mu}]^{\mathrm{T}}$ where ς, $\dot{\varsigma}$, μ, and $\dot{\mu}$ have the same same meaning as in Section 5.1. The state transition matrix in (92) is:

$$F = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The process noise covariance is $Q_{k-1} = I_4$, where I_4 denotes the 4×4 identity matrix. Furthermore, the measurement model is:

$$y_{k,s} = \begin{bmatrix} \sqrt{(\varsigma_k - \varsigma^s)^2 + (\mu_k - \mu^s)^2} \\ atan2\left(\varsigma_k, \mu_k\right) \end{bmatrix} + v_{k,s} \tag{99}$$

where: $s = 1, 2$ refers to S_1, S_2; $y_{k,s}$ denotes the measurement of radar s at time k; $[\varsigma^s, \mu^s]^T$ refers to the known position of radar s, specifically $[\varsigma^1, \mu^1]^T = [10, 10]^T$ and $[\varsigma^2, \mu^2]^T = [0, 0]^T$; $v_{k,s}$ is the heavy-tailed noise of radar s at time k generated as follows:

$$v_{k,1} \sim \begin{cases} \mathcal{N}(0, \Lambda_{k,1}), & \text{with probability } 0.9 \\ \mathcal{N}(0, 100\Lambda_{k,1}), & \text{with probability } 0.1 \end{cases} \tag{100}$$

$$v_{k,2} \sim \begin{cases} \mathcal{N}(0, \Lambda_{k,2}), & \text{with probability } 0.9 \\ \mathcal{N}(0, 100\Lambda_{k,2}), & \text{with probability } 0.1 \end{cases} \tag{101}$$

where $\Lambda_{k,1} = \begin{bmatrix} 15^2\,[\mathrm{m}^2] & 0 \\ 0 & 0.015^2\,[\mathrm{rad}^2] \end{bmatrix}$ and $\Lambda_{k,2} = \begin{bmatrix} 10^2\,[\mathrm{m}^2] & 0 \\ 0 & 0.01^2\,[\mathrm{rad}^2] \end{bmatrix}$. Initial values for x_0 and P_0 are set as follows:

$$x_0 = [50\,[\mathrm{m}], 10\,[\mathrm{m/s}], 80\,[\mathrm{m}], 20\,[\mathrm{m/s}]\,]^{\mathrm{T}} \tag{102}$$

$$P_0 = \begin{bmatrix} 1\,[\mathrm{m}^2] & 0 & 0 & 0 \\ 0 & 1\,[\mathrm{m}^2/\mathrm{s}^2] & 0 & 0 \\ 0 & 0 & 1\,[\mathrm{m}^2] & 0 \\ 0 & 0 & 0 & 1\,[\mathrm{m}^2/\mathrm{s}^2] \end{bmatrix}. \tag{103}$$

Monte Carlo simulations have been carried out with 100 independent trials. The CIFF of Section 4.1 and the proposed VBST-CIFF of Section 4.2 are compared in the above described multi-sensor target tracking scenario. Specifically, in CIFF, Q and R are fixed to $Q = Q_{k-1}$ for both sensors, $R = \Lambda_{k,1}$ for sensor 1, and $R = \Lambda_{k,2}$ for sensor 2. Conversely, in the proposed VBST-CIFF, $Q = Q_{k-1}$, while the scale matrix of each sensor as well as other parameters are adaptively estimated. The same values for the simulation time t, sampling interval T, forgetting factor τ, number of VB iterations N, as well as initial parameters of the scale matrix (δ, Δ), initial parameters of the DOF parameter (ϕ, Φ) of the previous single-sensor case-study of Section 5.1 have been adopted.

The results are presented in Figures 7–10 and Table 2. It is clear from Figures 7 and 8 that the proposed VBST-CIF provides, compared to conventional CIF, smaller RMSEs in both position and velocity for each local sensor during the whole simulation time. Figure 10 shows that the proposed VBST-CIFF outperforms the conventional CIFF, at the fusion center during the whole simulation time. Furthermore, it can be seen that in Table 2 the proposed VBST-CIFF provides an MRMSE improvement of 74% in position and 58% in velocity compared to conventional CIF with sensor 1, as well as an improvement of 62% in position and 42% in velocity compared to CIF with sensor 2. Moreover, the proposed VBST-CIFF provides an MRMSE improvement of 58% in position and 42% in velocity compared to CIFF.

Table 2. MRMSEs in multi-sensor target tracking.

Filter	Sensor	$\mathbf{RMSE}_p$	$\mathbf{RMSE}_v$
CIF	1	40.1874	7.5373
CIF	2	27.1914	5.4180
CIFF	Fusion	24.1697	5.4225
VBST-CIF	1	15.5891	3.5502
VBST-CIF	2	11.4030	3.2353
VBST-CIFF	Fusion	10.1193	3.1117

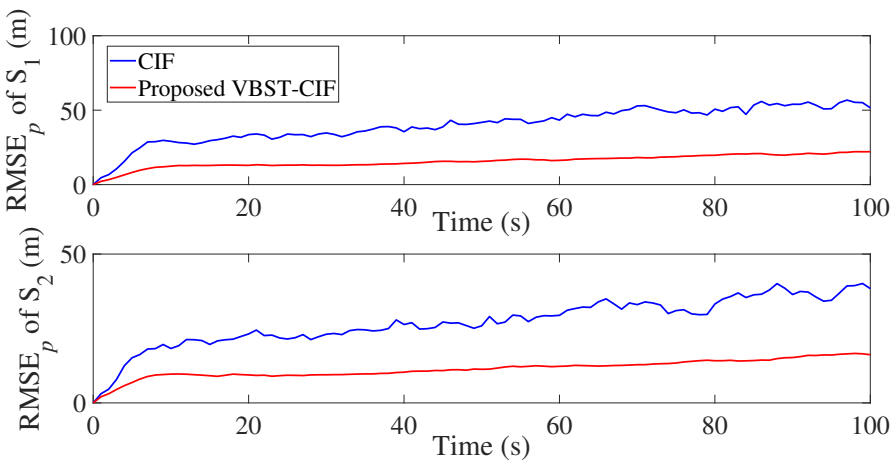

Figure 7. Position RMSE of different sensors in multi-sensor target tracking.

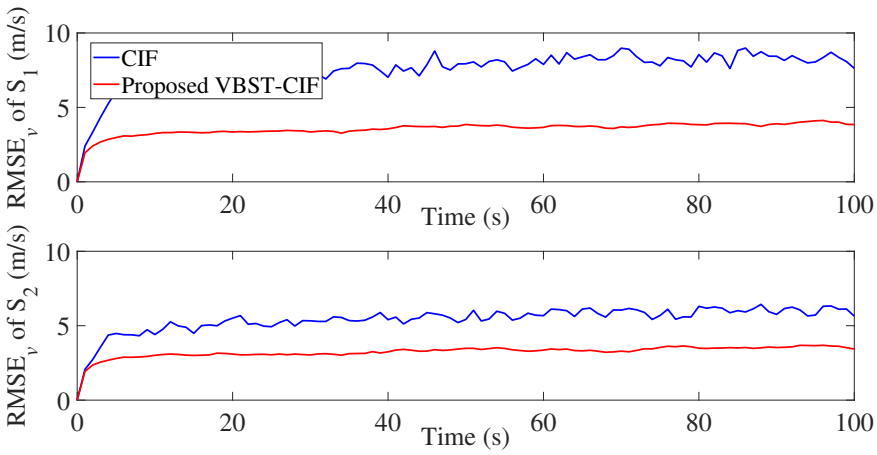

Figure 8. Velocity RMSE of different sensors in multi-sensor target tracking.

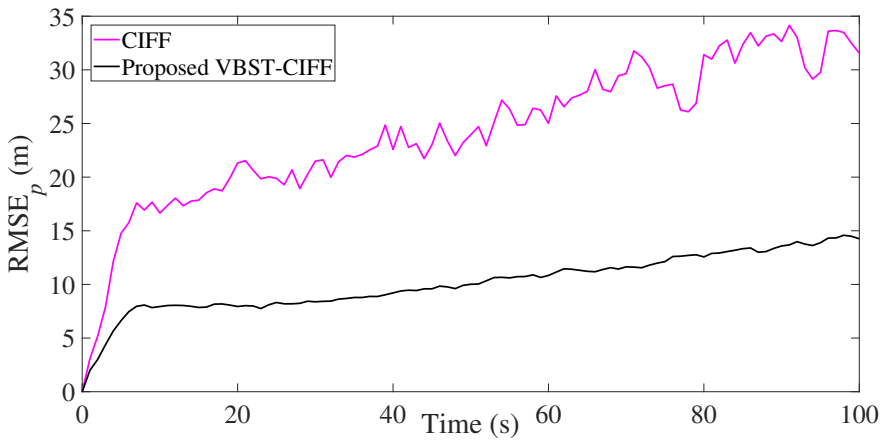

Figure 9. Position RMSE of fusion in multi-sensor target tracking.

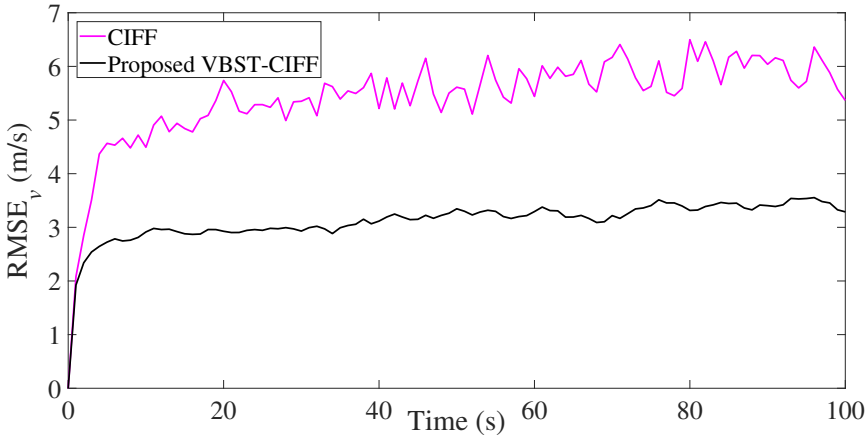

Figure 10. Velocity RMSE of fusion in multi-sensor target tracking.

5.3. Experimental Case-Study

Let us consider the experimental setup of Figure 11. A model-car moves along a straight line with a nearly constant velocity starting from initial position $(1018, 2619)$ [mm]. A depth camera (Camera 1, Intel RealSense Depth Camera D435i depth sensor) located at $(0, 0)$ measures its distance from the moving car, while another depth camera (Camera 2, Intel RealSense Tracking Camera T265), located onboard the car can get the location and azimuth angle of the car with respect to Camera 1. Please notice that the performance of Camera 1 is highly affected by the unstable Tiny YOLOv3 detector. Thus, the measurement function is the same as in (93) with a measurement noise of unknown heavy-tailed distribution. The total sampling samples are 363, with a sampling interval of $T = 1/60$ [s]. Moreover, initial values of the model-car are selected as:

$$x_0 = [1018\,[\mathrm{mm}],\ 2619\,[\mathrm{mm}],\ 228.76\,[\mathrm{mm/s}],\ 311.90\,[\mathrm{mm/s}]\,]^{\mathrm{T}} \tag{104}$$

$$P_0 = \begin{bmatrix} 1\,[\mathrm{mm}^2] & 0 & 0 & 0 \\ 0 & 1\,[\mathrm{mm}^2] & 0 & 0 \\ 0 & 0 & 1\,[\mathrm{mm}^2/\mathrm{s}^2] & 0 \\ 0 & 0 & 0 & 1\,[\mathrm{mm}^2/\mathrm{s}^2] \end{bmatrix}. \tag{105}$$

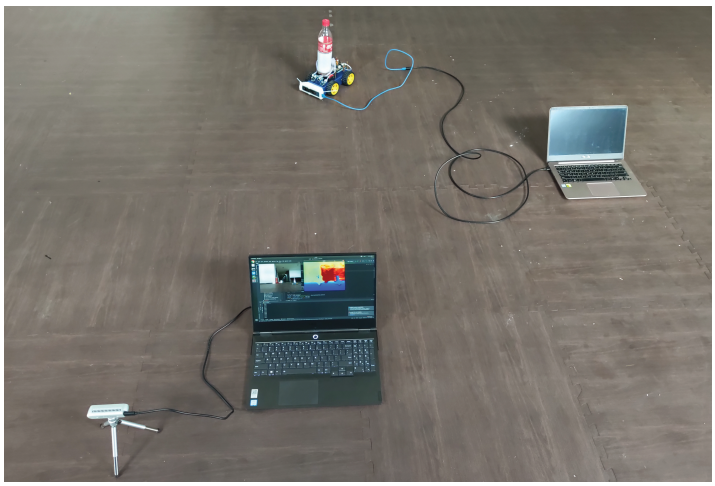

Figure 11. Experimental setup.

Let us consider the experimental results reported in Figures 12 and 13 and Table 3. Figure 12 shows the true and estimated car trajectories, from which it can be seen that the proposed VBST-CIF can track the car more accurately. Position and velocity RMSEs of both conventional CIF and proposed VBST-CIF are reported in Figure 13. Specifically, the proposed VBST-CIF provides an improvement of 46.24% in position MRMSE and of 0.08% in velocity MRMSE, with respect to conventional CIF, thus demonstratiung the effectiveness and superiority of the proposed VBST-CIF.

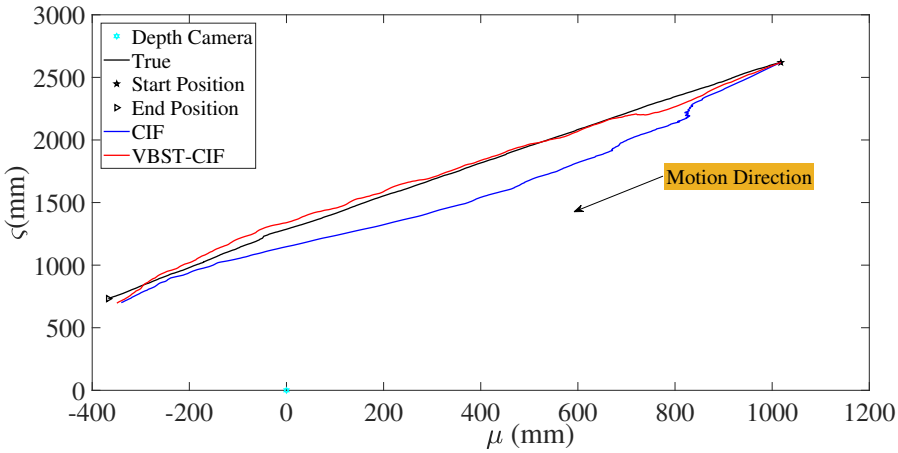

Figure 12. Car trajectory.

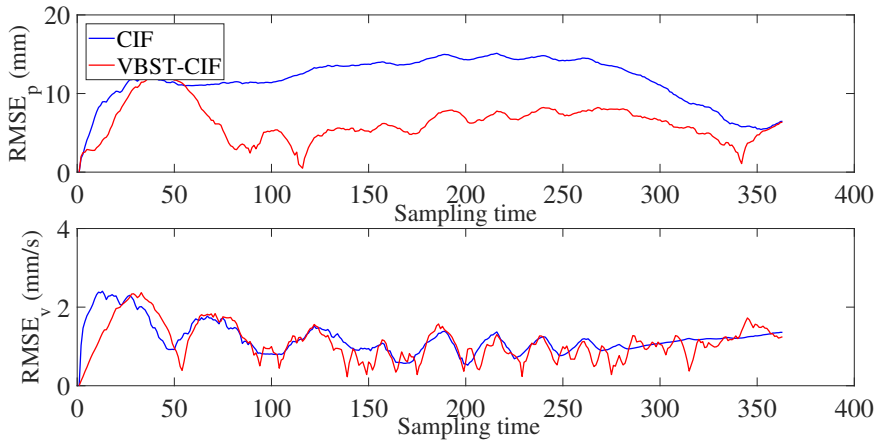

Figure 13. Position and velocity RMSEs.

Table 3. MRMSEs in the experimental case study.

Filter	RMSE_p [mm]	RMSE_v [mm/s]
CIF	11.7397	1.1985
VBST-CIF	6.3111	1.0991

6. Conclusions

This paper focused on state estimation and information fusion in multi-sensor nonlinear systems with heavy-tailed measurement noise. An adaptive variational Bayes Student's t cubature information filter (VBST-CIF) algorithm was first designed for a nonlinear single-sensor system based on the cubature information filter (CIF) framework, wherein the Student's t (ST) distribution was utilized to model the heavy-tailed measurement noise, and the variational Bayes (VB) method along with the spherical-radial cubature (SCR) rule were exploited in order to jointly estimate the system state as well as noise statistics. The VBST-CIF facilitated multi-sensor fusion, allowing to derive

a VBST cubature information feedback fusion (VBST-CIFF) algorithm for nonlinear multi-sensor systems with heavy-tailed measurement noise. The filtering accuracy of the proposed algorithms was assessed in single-sensor and multi-sensor target tracking scenarios. Simulation and experimental results demonstrated that the proposed VBST-CIF/VBST-CIFF outperformed conventional CIF/CIFF algorithms. Future work will focus on consensus filtering for multi-agent systems with heavy-tailed process and measurement noise.

Author Contributions: X.D.: Conceptualization, methodology, and writing-original draft; L.C.: writing—review & editing, supervision; Y.C.: writing—review & editing, funding acquisition, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by National Natural Science Foundation of China (61627810), National Science and Technology Major Program of China (2018YFB1305003).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kalman, R.E. A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.* **1960**, *82*, 34-45. [CrossRef]
2. Tsyganova, J.V.; Kulikova, M.V. SVD-based Kalman filter derivative computation. *IEEE Trans. Autom. Control* **2017**, *62*, 4869–4875. [CrossRef]
3. Lu, X.; Wang, L.; Wang, H.; Wang, X. Kalman filtering for delayed singular systems with multiplicative noise. *IEEE/CAA J. Autom. Sin.* **2016**, *3*, 51–58.
4. Elliott, R.J. Filtering with uncertain noise. *IEEE Trans. Autom. Control* **2017**, *62*, 876–881. [CrossRef]
5. Nurminen, H.; Ardeshiri, T.; Piché, R.; Gustafsson, F. Robust inference for state-space models with skewed measurement noise. *IEEE Signal Process. Lett.* **2015**, *22*, 1898–1902. [CrossRef]
6. Liu, Q.; Luo, Z.; Huang, J.; Yan, S. A comparative study of four kinds of adaptive decomposition algorithms and their applications. *Sensors* **2018**, *18*, 2120. [CrossRef] [PubMed]
7. Fan, Y.; Zhang, Y.; Wang, G.; Wang, X.; Li, N. Maximum correntropy based unscented particle filter for cooperative navigation with heavy-tailed measurement noises. *Sensors* **2018**, *18*, 10. [CrossRef]
8. Liu, Z.; Chen, S.; Wu, H.; He, R.; Hao, L. A Student's t mixture probability hypothesis density filter for multi-target tracking with outliers. *Sensors* **2018**, *18*, 4. [CrossRef]
9. Zhang, X.; Ding, F.; Xu, L.; Yang, E. Highly computationally efficient state filter based on the delta operator. *Int. J. Adapt. Control Signal Process.* **2019**, *33*, 875–889. [CrossRef]
10. Zhang, X.; Ding, F.; Yang, E. State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors. *Int. J. Adapt. Control Signal Process.* **2019**, *33*, 1157–1173. [CrossRef]
11. Zhang, X.; Ding, F. Recursive parameter estimation and its convergence for bilinear systems. *IET Control Theory Appl.* **2020**, *14*, 677–688. [CrossRef]
12. Zhang, X.; Ding, F. Adaptive parameter estimation for a general dynamical system with unknown states. *Int. J. Robust Nonlinear Control* **2020**, *30*, 1351–1372. [CrossRef]
13. Ding, F. Combined state and least squares parameter estimation algorithms for dynamic systems. *Appl. Math. Model.* **2014**, *38*, 403–412. [CrossRef]
14. Ding, F. State filtering and parameter estimation for state space systems with scarce measurements. *Signal Process.* **2014**, *104*, 369-380. [CrossRef]
15. Cao, Y.; Lu, H.; Wen, T. A safety computer system based on multi-sensor data processing. *Sensors* **2019**, *19*, 818. [CrossRef] [PubMed]
16. Jin, X.B.; Yang, N.X.; Wang, X.Y.; Bai, Y.T.; Kong, J.L. Hybrid deep learning predictor for smart agriculture sensing based on empirical mode decomposition and gated recurrent unit group model. *Sensors* **2020**, *20*, 1334. [CrossRef] [PubMed]
17. Bai, Y.T.; Wang, X.Y.; Jin, X.B.; Zhao, Z.Y.; Zhang, B.H. A neuron-based Kalman filter with nonlinear autoregressive model. *Sensors* **2020**, *20*, 299. [CrossRef]
18. Wang, J.; Wang, J.; Zhang, D.; Shao, X. Stochastic feedback based Kalman filter for nonlinear continuous-discrete systems. *IEEE Trans. Autom. Control* **2018**, *63*, 3002–3009. [CrossRef]

19. Athans, M.; Wishner, R.; Bertolini, A. Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements. *IEEE Trans. Autom. Control* **1968**, *13*, 504–514. [CrossRef]
20. Vemula, M.; Bugallo, M.F.; Djuric, P.M. Performance comparison of Gaussian-based filters using information measures. *IEEE Signal Process. Lett.* **2007**, *14*, 1020–1023. [CrossRef]
21. Wang, X.; Yaz, E.E. Second-order fault tolerant extended Kalman filter for discrete time nonlinear systems. *IEEE Trans. Autom. Control* **2019**, *64*, 5086–5093. [CrossRef]
22. Liu, Y.; Wang, H.; Hou, C. UKF based nonlinear filtering using minimum entropy criterion. *IEEE Trans. Signal Process.* **2013**, *61*, 4988–4999. [CrossRef]
23. Arasaratnam, I.; Haykin, S. Cubature Kalman filters. *IEEE Trans. Autom. Control* **2009**, *54*, 1254–1269. [CrossRef]
24. Arasaratnam, I.; Haykin, S.; Hurd, T.R. Cubature Kalman filtering for continuous-discrete systems: Theory and simulations. *IEEE Trans. Signal Process.* **2010**, *58*, 4977–4993. [CrossRef]
25. Arasaratnam, I. Sensor fusion with square-root cubature information filtering. *Intell. Control Autom.* **2013**, *4*, 11–17. [CrossRef]
26. Ge, Q.; Shao, T.; Yang, Q.; Shen, X.; Wen, C. Multisensor nonlinear fusion methods based on adaptive ensemble fifth-degree iterated cubature information filter for biomechatronics. *IEEE Trans. Syst. Man Cybern. Syst.* **2016**, *46*, 912–925. [CrossRef]
27. Dong, P.; Jing, Z.; Leung, H.; Shen, K. Variational Bayesian adaptive cubature information filter based on Wishart distribution. *IEEE Trans. Autom. Control* **2017**, *62*, 6051–6057. [CrossRef]
28. Piché, R.; Särkkä, S.; Hartikainen, J. Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate Student-t distribution. In Proceedings of the 2012 IEEE International Workshop on Machine Learning for Signal Processing, Santander, Spain, 23–26 September 2012; pp. 1–6.
29. Lee, K.; Johnson, E.N. Robust outlier-adaptive filtering for vision-aided inertial navigation. *Sensors* **2020**, *20*, 2036. [CrossRef]
30. Barkat, B.; Yingtuo, J.; Abed-Meraim, K. Detection of known FM signals in known heavy-tailed noise. In Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology, Darmstadt, Germany, 17 December 2003; pp. 130–133.
31. Roth, M.; Özkan, E.; Gustafsson, F. A Student's t filter for heavy tailed process and measurement noise. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, 26–31 May 2013; pp. 5770–5774.
32. Kapoor, R.; Tsihrintzis, G.A.; Nandhakumar, N. Detection of obscured targets in heavy-tailed radar clutter using an ultra-wideband (UWB) radar and alpha-stable clutter models. In Proceedings of the Conference Record of The Thirtieth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 3–6 November 1996; Volume 2, pp. 863–867.
33. Huang, Y.; Zhang, Y.; Li, N.; Chambers, J. A robust Gaussian approximate filter for nonlinear systems with heavy tailed measurement noises. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 4209–4213.
34. Taghavi, E.; Tharmarasa, R.; Kirubarajan, T.; Mcdonald, M. Multisensor-multitarget bearing-only sensor registration. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 1654–1666. [CrossRef]
35. Li, X.; Jilkov, V. Survey of maneuvering target tracking. Part V. multiple-model methods. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1255–1321.
36. Pan, Q.; Hu, Y.; Lan, H.; Sun, S.; Wang, Z.; Yang, F. Information fusion progress: Joint optimization based on variational Bayesian theory. *Zidonghua Xuebao/Acta Autom. Sin.* **2019**, *45*, 1207–1223.
37. Wang, X.; Cui, N.; Guo, J. Huber-based unscented filtering and its application to vision-based relative navigation. *IET Radar Sonar Nav.* **2010**, *4*, 134–141. [CrossRef]
38. Huang, Y.; Zhang, Y.; Li, N.; Wu, Z.; Chambers, J.A. A novel robust Student's t-based Kalman filter. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 1545–1554. [CrossRef]
39. Agamennoni, G.; Nieto, J.I.; Nebot, E.M. An outlier-robust Kalman filter. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 1551–1558.
40. Agamennoni, G.; Nieto, J.I.; Nebot, E.M. Approximate inference in state-space models with heavy-tailed noise. *IEEE Trans. Signal Process.* **2012**, *60*, 5024–5037. [CrossRef]

41. Huang, Y.; Zhang, Y.; Li, N.; Chambers, J. A robust Gaussian approximate fixed-interval smoother for nonlinear systems with heavy-tailed process and measurement noises. *IEEE Signal Process. Lett.* **2016**, *23*, 468–472. [CrossRef]
42. Huang, Y.; Zhang, Y.; Li, N.; Naqvi, M.; Chambers, J. A robust Student's t based cubature filter. In Proceedings of the 2016 19th International Conference on Information Fusion (FUSION), Heidelberg, Germany, 5–8 July 2016; pp. 9–16.
43. Chandra, K.; Gu, D.; Postlethwaite, I. Square root cubature information filter. *IEEE Sens. J.* **2013**, *13*, 750–758. [CrossRef]
44. Huang, Y.; Zhang, Y.; Wu, Z.; Li, N.; Chambers, J. A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices. *IEEE Trans. Autom. Control* **2018**, *63*, 594–601. [CrossRef]