Distributed Consensus Student-*t* Filter for Sensor Networks With Heavy-Tailed Process and Measurement Noises

JINRAN WANG[1], PENG DONG[ID,2], KAI SHEN[ID,3], XUN SONG[1], AND XIAODONG WANG[1]

[1]State Key Laboratory of Intelligent Manufacturing System Technology, Beijing Institute of Electronic System Engineering, Beijing 100854, China
[2]School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, China
[3]School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China

Corresponding author: Peng Dong (dongpengkty@sjtu.edu.cn)

ABSTRACT In the estimation of distributed sensor networks, process noise and measurement noise may have outliers which have heavy-tailed characteristics. To solve this problem, this paper proposes a distributed consensus estimating method for sensor networks based on Student-*t* distribution. In the state space model, both process noise and measurement noise are modeled as Student-*t* distributions with heavy-tailed characteristics. First, for the assumption that the process noise and measurement noise have the same degree of freedom parameters, an exact distributed consensus Student-*t* filtering algorithm is derived. In practical applications, this assumption is often not true, and due to the increasing degrees of freedom, the method will quickly converge to the traditional distributed consensus Kalman filter. Therefore, it is necessary to relax the assumption of the same degree of freedom and keep the degree of freedom unchanged within a certain range. Based on this, an approximate distributed consensus Student-*t* filter algorithm is proposed. Simulation results verify the effectiveness of the proposed algorithm.

INDEX TERMS Student-*t* distribution, distributed consensus filter, distributed sensor networks.

NOMENCLATURE

$N(\cdot; m, (\Lambda)^{-1})$	Gaussian distribution with mean m and precision matrix Λ
$(\cdot)^T$	Matrix transpose manipulation
$(\cdot)^{-1}$	Matrix inversion operation
ℓ	Iteration index
$\Gamma(\cdot)$	Gamma function
$\hat{x}_{k\|k-1}$	Prior estimate of x_k
$\mathcal{A}$	Set of connections between nodes
$\mathcal{N}^i$	Neighbors of node i
$\mathcal{N}$	Set of sensor nodes
$\mathcal{Z}_k$	Measurements of all sensor nodes till time k
$\mathscr{Z}_k$	measurements of all sensor nodes at time k
$\|\cdot\|$	Cardinality of a set
Ω^i	Information matrix
$\pi^{i,j}$	Consensus weight
$\tilde{z}_k^i$	Residual vector
d_z^i	Dimension of measurement vector
F_k	State transition matrix
H_k^i	Measurement matrix
i	Sensor index
k	Time instant index
K_k	Filter gain
$KL(\cdot)$	Kullback-Leibler divergence
L	Conseunsu steps
n_x	Dimension of state vector
$p(\cdot)$	Probability density function
q^i	Information vectors
Q_k	Scale matrix for w_k
R_k^i	Scale matrix for v_k^i
x_k	System state vector
Z_k^i	Measurements of the node i till time k
$\hat{x}_k$	Posterior estimate of x_k
$G(\cdot; a, b)$	Gamma distribution with shape parameter a and scale parameter b

$\text{St}(\cdot; m, P, \nu)$	Student-*t* distribution with mean m, scale matrix P and degree of freedom ν
H_k	$[H_k^1, H_k^2, \ldots, H_k^i]^T$
P_k	Scale matrix for x_k
R_k	$diag(R_k^1, R_k^2, \ldots, R_k^i)$
v_k	$[(v_k^1)^T, (v_k^2)^T, \ldots, (v_k^i)^T, \ldots]^T$
v_k^i	Measurement noise
w_k	Process noise
z_k^i	Measurement vector

I. INTRODUCTION

Distributed state estimation is important in distributed sensor networks [1], [2]. Due to the complex environment, the communication, perception and processing capabilities of the distributed sensor network will be limited. Thus, traditional data sharing and fusion methods are not applicable in this case. There are three major methods for distributed state estimation, such as consensus-based algorithms [3], gossip-based algorithms [4] and diffusion-based algorithms [5]. Consensus-based algorithms have capability to provide better performance in terms of estimate accuracy. Information sharing and interaction using the consensus method requires neither a fusion center nor a full connection between network nodes. Information interaction only occurs between neighboring nodes, and the information of all nodes can eventually be coordinated and consistent. This method is applicable to any network topology and can simultaneously improve the flexibility and robustness of the network.

The combination of consensus theory and Kalman filter is the most direct idea of consensus estimation. This combination can apply the consensus mechanism to the prediction or update step of Kalman filter, so as not to lose the basic characteristics of the Kalman filter. The consensus on estimation (CE) methods using consensus strategy for the state estimation of each node are proposed in [6]. To teal with the conservative characteristic of the estimation for the CE method, a consensus on measurement (CM) method which makes the local new information pair to reach consensus is proposed in [7]. However, this method needs large number of consistent steps to achieve convergence. A consensus on information (CI) method utilizing uniform local averaging of information matrices and information vectors is proposed in [8]. Many scholars have conducted in-depth research under this framework and achieved fruitful results [9]–[17].

Most of the consensus filters assume that the process and measurement noises are Gaussian distributions. However, in many practical situations, the process and measurement noises may suffer from outliers, which may come from unreliable sensors, unmodeled anomalies, sudden disturbance in the system environment, or target maneuvers [18], [19]. The Gaussian distribution assumption may cause poor performance or system failures in these situations. Thus the Student-*t* distribution with heavy-tailed is used to model the uncertainties exhibiting frequent occurrence of the outliers [20]–[24]. A linear distributed consensus filter with CI

strategy to handle measurement outliers is proposed in [25], where the measurement noise of each sensor node is modeled by the multivariate Student-*t* distribution and variational Bayesian (VB) method [26], [27] is used to approximate the joint posterior density. Then it is extended to nonlinear cases in [28] where hybrid consensus strategy [9], [11] is used. However, these methods can only handle the scenarios with heavy-tailed measurement noise and well-behaved process noise. The particle filter can handle the process and measurement noises with arbitrary distributions [29], however, it suffers from curse of dimensionality in high dimensional problems. Gaussian sum filter (GSF) [30] can deal with heavy-tailed non-Gaussian noises but it needs a lot of Gaussian distributions to model the heavy-tailed process and measurement noises accurately. There are other robust methods such as WLAV (weighted least absolute value) filter [31] and MEAV (Maximum Exponential Absolute Value) filter [32] to deal with bad data. However, these methods need optimization or iteration procedure which may lead to a lot of extra computation. Recently, the robust Student-*t* filters for heavy-tailed process and measurement noises have been proposed in [19], [33]–[35] for single sensor. This method has less computational complexity, is easy to apply and can deal with high-dimensional problems.

A distributed consensus Student-*t* filter is presented in this paper to handle both heavy-tailed process and measurement noises for distributed sensor networks. The main contributions of this paper can be highlighted as follows.

1) Both state and measurement of each sensor node are modeled by Student-*t* distribution. Under certain assumptions, the distributed consensus Student-*t* filter is derived: the recursion predicted and updates steps for multiple sensors are derived first, then the CI strategy for Student-*t* distribution are derived based on moment matching.

2) Since the assumptions are too restricted for real scenario, some approximation are made to relax the assumptions for practical applications. In addition, the degree of freedom for Student-*t* distribution does not grow over time to maintain heavy-tailed characteristics.

3) The simulation are processed in scenario where process noise and measurement noise are heavy-tailed. The results show that the proposed method outperforms the conventional distributed consensus filter.

The remainder of this paper is organized as follows. Section II describes models of the sensor network and the consensus method used in this paper. Section III presents our distributed consensus Student-*t* filter. Simulation results and analysis are given in Section IV and the conclusion is given in Section V.

II. PROBLEM FORMALIZATION

In this paper, we consider the sensor network represented by $(\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ denotes set of sensor nodes and $\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of connections between nodes such

that $(i, j) \in \mathcal{A}$ if node j can receive data from node i. The set $\mathcal{N}^i \triangleq \{j : (j, i) \in \mathcal{A}\}$ denotes neighbors (including i itself) of node $i \in \mathcal{N}$.

Consider a discrete time linear system

$$x_k = F_{k-1} x_{k-1} + w_{k-1}, \tag{1}$$

and measurement equations of the sensor node $i \in \mathcal{N}$

$$z_k^i = H_k^i x_k + v_k^i, \quad i \in \mathcal{N}, \tag{2}$$

where x_k is the n_x-dimension state vector, F_k is the state transition matrix, w_k is the process noise, z_k^i is the d_z^i-dimension measurement vector of node i, H_k^i is the measurement matrix of node i, v_k^i is the measurement noise of node i.

For the initial state x_0, the process noise w_k and the measurement noise v_k^i, we have the following assumptions. Assume that the initial state and the noise signal are not related to each other, and their marginal distributions are

$$p(x_0) = \text{St}(x_0; \hat{x}_0, P_0, \eta_0), \tag{3}$$

$$p(w_k) = \text{St}(w_k; 0, Q_k, \gamma_k), \tag{4}$$

$$p(v_k^i) = \text{St}(v_k^i; 0, R_k^i, \delta_k^i), \tag{5}$$

where $\text{St}(\cdot; m, P, \nu)$ denotes Student-*t* distribution with mean m, scale matrix P and degree of freedom (DOF) ν. Thus, η_0, γ_k and δ_k^i are the DOF of related densities. P_0, Q_k and R_k^i are the scale matrix of related densities.

It is difficult to obtain a globally closed solution by maximizing the likelihood of Student-t distribution, however, we can decompose Student-t distribution into a mixture of infinite Gaussian distributions, which have the same mean and precision. The Student-t distribution after decomposition can be expressed as follows:

$$\text{St}(x; m, P, \nu) = \int_0^\infty \text{N}(x; m, (u\nu)^{-1})$$
$$\times \text{G}(u; \nu/2, \nu/2) du, \tag{6}$$

where $N(x; m, (\Lambda)^{-1})$ is Gaussian distribution with mean m and precision matrix Λ, $\text{G}(u; a, b)$ is Gamma distribution with shape parameter a and scale parameter b. The probability density function (PDF) of Student-*t* distribution is given by

$$\text{St}(x; m, P, \nu) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{(\nu\pi)^{\frac{d}{2}}} \frac{1}{\sqrt{\det(P)}}$$
$$\times \left(1 + \frac{1}{\nu}(x-m)^T P^{-1}(x-m)\right)^{-\frac{\nu+d}{2}} \tag{7}$$

where $\Gamma(\cdot)$ denotes the Gamma function. It should be noted that P is not a covariance matrix in general while $\frac{\nu}{\nu-2}P$ is the covariance for $\nu > 2$. When $\nu = 1$, it converges to the Cauchy distribution, while $\nu \to \infty$ it becomes the Gaussian distribution. Compared with the Gaussian distribution, the Student-*t* distribution has the heavy tails.

Consensus algorithm is the information exchange rule that ensures that the amount of concern of each node achieves consistency. The weighted Kullback-Leibler average $\bar{p}(\cdot)$

among the probability density function (PDF) $\{p^i(\cdot)\}$ are given by [11]

$$\bar{p} = \arg\inf_{p(\cdot)} \sum_{i \in \mathcal{N}} \pi^i KL(p||p^i) \tag{8}$$

where $\pi^i > 0$ is the weight and $\sum_{i \in \mathcal{N}} \pi^i = 1$, $KL(p||p^i)$ is the Kullback-Leibler divergence between the PDFs $p(\cdot)$ and $p^i(\cdot)$. The problem of probability density consensus can be treated as finding a consensus algorithm to make

$$\lim_{l \to +\infty} p_l^i = \bar{p}(x), \quad \forall i \in \mathcal{N} \tag{9}$$

where $\bar{p}(\cdot)$ is the asymptotic PDF. Thus the solution to (8) is

$$\bar{p}(x) = \frac{\prod_{i \in \mathcal{N}} [p^i(x)]^{\pi^i}}{\int \prod_{i \in \mathcal{N}} [p^i(x)]^{\pi^i} dx} \triangleq \bigoplus_{i \in \mathcal{N}} (\pi^i \odot p^i(x)) \tag{10}$$

where $\pi^i = 1/|\mathcal{N}|$, the operators $\oplus$ and $\odot$ are given by

$$p(x) \oplus q(x) \triangleq \frac{p(x)q(x)}{\int p(x)q(x)dx} \tag{11}$$

$$\pi \odot p(x) \triangleq \frac{[p(x)]^\pi}{\int [p(x)]^\pi dx} \tag{12}$$

Therefore, the solution can be obtained by exchanging the local data with the neighbors via convex combination in a iterative way

$$p_\ell^i(x) = \bigoplus_{j \in \mathcal{N}^i} (\pi^{i,j} \odot p_{\ell-1}^j(x)) \tag{13}$$

where $\pi^{i,j} \geq 0$ is the consensus weight and $\sum_{j \in \mathcal{N}^i} \pi^{i,j} = 1$, ℓ is the iteration index and iterations are initialized with $p_0^i(x) = p^i(x)$.

The purpose of consensus filter in this paper is to obtain a consensus state of a sensor network with both heavy-tailed process noise and measurement noise.

III. PROPOSED METHODS

A. THE EXACT DISTRIBUTED CONSENSUS STUDENT-*t* FILTER

1) THE STUDENT-*t* FILTER FOR MULTIPLE SENSORS

Suppose $Z_k^i = \{z_1^i, z_2^i, \ldots, z_k^i, i \in \mathcal{N}\}$ denotes measurements of the node i till time k, $\mathcal{Z}_k = \{Z_k^i, i \in \mathcal{N}\}$ denotes measurements of all sensor nodes till time k and $\mathcal{Z}_k = \{z_k^i, i \in \mathcal{N}\}$ denotes measurements of all sensor nodes at time k. Similar to the distributed consensus Kalman filter, we divide the filter recursion into time update and measurement update.

Suppose DOF $\gamma_k = \eta_k$ for all process noises, where η_k is the DOF of $x_k | \mathcal{Z}_k$, then we have

$$p(x_{k-1}, w_{k-1} | \mathcal{Z}_{k-1})$$
$$= \text{St}\left(\begin{bmatrix} x_{k-1} \\ w_{k-1} \end{bmatrix}; \begin{bmatrix} \hat{x}_{k-1} \\ 0 \end{bmatrix}, \begin{bmatrix} P_{k-1} & 0 \\ 0 & Q_{k-1} \end{bmatrix}, \eta_{k-1}\right) \tag{14}$$

The predicted density is

$$p(x_k | \mathcal{Z}_{k-1}) = \text{St}(x_k; \hat{x}_{k|k-1}, P_{k|k-1}, \eta_{k-1}) \tag{15}$$

Thus the parameter η_k is not changed.

We assume that all nodes have the same DOF in measurement noise distribution $\delta_k^i = \delta_k$ and $\delta_k = \eta_{k-1}$, then we can obtain the joint density of the predicted state and the measurement noise

$$p(x_k, v_k | \mathcal{Z}_{k-1})$$
$$= \text{St}\left(\begin{bmatrix} x_k \\ v_k \end{bmatrix}; \begin{bmatrix} \hat{x}_{k|k-1} \\ 0 \end{bmatrix}, \begin{bmatrix} P_{k|k-1} & 0 \\ 0 & R_k \end{bmatrix}, \eta_{k-1} \right) \quad (16)$$

where

$$v_k = [(v_k^1)^T, (v_k^2)^T, \dots, (v_k^i)^T, \dots]^T, \quad i \in \mathcal{N} \quad (17)$$
$$R_k = diag(R_k^1, R_k^2, \dots, R_k^i), \quad i \in \mathcal{N} \quad (18)$$

Then, the joint density of state and measurement can be obtained by linear transformation

$$p(x_k, z_k | \mathcal{Z}_{k-1}) = \text{St}\left(\begin{bmatrix} x_k \\ \mathcal{Z}_k \end{bmatrix}; \begin{bmatrix} \hat{x}_{k|k-1} \\ H_k \hat{x}_{k|k-1} \end{bmatrix}, \right.$$
$$\left. \begin{bmatrix} P_{k|k-1} & P_{k|k-1} H_k^T \\ H_k P_{k|k-1} & S_k \end{bmatrix}, \eta_{k-1} \right) \quad (19)$$

where

$$S_k = H_k P_{k|k-1} H_k^T + R_k \quad (20)$$
$$H_k = [H_k^1, H_k^2, \dots, H_k^i], \quad i \in \mathcal{N} \quad (21)$$

Thus, given all measurement $\mathcal{Z}_k$, the conditional density of the state is still Student-*t* distribution

$$p(x_k | \mathcal{Z}_k) = \text{St}(x_k; \hat{x}_k, P_k, \eta_k) \quad (22)$$

The above is the recursive process of obtaining all sensor measurements.

2) CONSENSUS FOR STUDENT-*t* DISTRIBUTION

If the PDF of each sensor node is Gaussian distribution such as $p^i(x) = N(x^i, (\Lambda^i)^{-1})$, then we can obtain the average PDF by averaging information vectors $q^i = \Omega^i \hat{x}^i$ and information matrix $\Omega^i = \Lambda^i$, and the following consensus algorithm are derived in [10]

$$\Omega_\ell^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} \Omega_{k,\ell-1}^j \quad (23)$$

$$q_\ell^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} q_{\ell-1}^j \quad (24)$$

For the Student-*t* distribution, the exact form of consensus algorithm can hardly be obtained. It can be noted that the Student-*t* distribution $\text{St}(x; m, P, \nu)$ will converge to the Gaussian distribution as $\nu \to \infty$. Therefore, we can approximate the PDF $p(x) = \text{St}(x; m, P_k, \nu)$ by a Gaussian distribution $p'(x) = N(x; m, \bar{P}) \approx \text{St}(x; m, \bar{P}, \bar{\nu})$ to take the advantage of the Gaussian consensus algorithm such as (23) and (24). That is

$$\text{St}(x; m, P, \nu) \approx \text{St}(x; m, \bar{P}, \bar{\nu}) \quad (25)$$

where DOF $\bar{\nu} \to \infty$.

Qualitative features should be retained in the problem of adjusting the matrix parameters given a new DOF. Therefore,

the adjusted matrix parameters $\bar{P}$ should be scaled versions of the original matrix given by $\bar{P} = cP$. As a general problem, we must find a scalar $c > 0$ so that the PDF $p(x)$ and $p'(x)$ is close in some respects. Once $\bar{\nu}$ is given, the parameter c can be found using moment matching method. According to moment matching method, we obtain the condition

$$\frac{\nu}{\nu - 2} P = \frac{\bar{\nu}}{\bar{\nu} - 2} cP \quad (26)$$

for $\nu > 2$ and $\bar{\nu} \to \infty$, then the scale factor is given by

$$c = \frac{\nu(\bar{\nu} - 2)}{(\nu - 2)\bar{\nu}} = \frac{\nu}{\nu - 2} \quad (27)$$

According to the approximation, we can use the consensus steps (23) and (24) directly. After the the consensus steps, we should do some inverse operations to change the DOF back to ν, and specific steps are given in Sec. III-A3.

3) THE RECURSION OF EXACT DISTRIBUTED CONSENSUS STUDENT-*t* FILTER

According to the predicted and update steps in Sec. III-A1, and the consensus step in Sec. III-A2, then for each local sensor in the distributed sensor network, the following exact consensus Student-*t* filter recursive process can be obtained:

(1) Prediction of local filter:

$$\hat{x}_{k|k-1}^i = F_k \hat{x}_{k-1}^i \quad (28)$$
$$P_{k|k-1}^i = F_k P_{k-1}^i F_k^T + Q_{k-1} \quad (29)$$

(2) Update of local filter:

$$\tilde{z}_k^i = z_k^i - H_k^i \hat{x}_{k|k-1}^i \quad (30)$$
$$S_k^i = H_k^i P_{k|k-1}^i (H_k^i)^T + R_k^i \quad (31)$$
$$K_k^i = P_{k|k-1}^i (H_k^i)^T / S_k^i \quad (32)$$
$$x_{k,0}^i = \hat{x}_{k|k-1}^i + K_k^i \tilde{z}_k^i \quad (33)$$
$$P_{k,0}^i = \frac{\eta_{k-1}^i + \Delta_{z,k}^i}{\eta_{k-1}^i + d_z^i} (P_{k|k-1}^i - K_k^i S_k^i (K_k^i)^T) \quad (34)$$
$$\Delta_{z,k}^i = (\tilde{z}_k^i)^T (S_k^i)^{-1} \tilde{z}_k^i \quad (35)$$
$$\eta_k^i = \eta_{k-1}^i + d_z^i \quad (36)$$

(3) Consensus on the information matrix and information vector: Approximate the initial information matrix and information vector by

$$\bar{P}_{k,0}^i = \frac{\eta_k^i}{\eta_k^i - 2} P_{k,0}^i \quad (37)$$
$$\Omega_{k,0}^i = (\bar{P}_{k,0}^i)^{-1} \quad (38)$$
$$q_{k,0}^i = \Omega_{k,0}^i \hat{x}_{k,0}^i \quad (39)$$

For a L-step consensus iteration, the consensus on posterior information is carried out in the form

$$\Omega_{k,\ell}^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} \Omega_{k,\ell-1}^j, \quad \ell = 1, \dots, L \quad (40)$$

$$q_{k,\ell}^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} q_{k,\ell-1}^j, \quad \ell = 1, \dots, L \quad (41)$$

where $\ell = 1, 2, \cdots, L$ is the consensus step, $\pi^{i,j}$ is the consensus weight. A convex combination is adopted by supposing $\pi^{i,j} \geq 0$ and $\sum_{j \in \mathcal{N}^i} \pi^{i,j} = 1$.

(4) Recover the state and the scale matrix:

$$\hat{x}_k^i = (\Omega_{k,L}^i)^{-1} q_{k,L}^i \tag{42}$$

$$P_k^i = \frac{\eta_k^i - 2}{\eta_k^i} (\Omega_{k,L}^i)^{-1} \tag{43}$$

It can be noted that with each measurement update, the degrees of freedom increase according to (36). In turn, this requires an increase in the degree of freedom of noise, making the problem more and more Gaussian.

B. THE APPROXIMATED DISTRIBUTED CONSENSUS STUDENT-*t* FILTER

The conditions required in Section 3.1 will hardly be met in practice. Therefore, we introduce some approximations, and the resulting filtering algorithm is only slightly more complicated than the exact filter in (28)-(34). In addition, we also prevent an increase in degrees of freedom, thereby maintaining heavy-tailed density for the entire time.

Under more practical assumptions, we consider the linear models (1) and (2) again. At this time, the degrees of freedom γ_k and δ_k in equations (4) and (5) are arbitrary, therefore the closed form of time update and measurement update cannot be derived in closed form. Suppose at time k we have the following posterior density

$$p(x_k | \mathcal{Z}_k) = \mathrm{St}(x_k; \hat{x}_k, P_k, \eta_k) \tag{44}$$

In the process of deducing the exact filter, we use a formula to express the joint density of x_k and w_k after Z_k is given. If $\gamma_k \neq \eta_k$, then $p(x_k, w_k | \mathcal{Z}_k)$ will not be represented as a closed form unless w_k is independent. However, for independent noise, the joint density

$$p(x_k, w_k | \mathcal{Z}_k) = \mathrm{St}(x_k; \hat{x}_k, P_k, \eta_k) \mathrm{St}(w_k; 0, Q_k, \gamma_k) \tag{45}$$

is no longer Student-*t* distribution nor ellipsoid shape. Therefore, we cannot derive a concise time update equation. What leads to the convenient expression in the exact filter (28)-(34) is the joint density in (15) and (16). Therefore, it is a reasonable choice to force the actual state and noise density into this format.

For time update, we suggest to find the common degree of freedom parameter $\tilde{\eta}_k$ from γ_k to η_k. Similar to Sec. 3.1, we use

$$p(x_{k-1}, w_{k-1} | \mathcal{Z}_{k-1})$$
$$\approx \mathrm{St}\left(\begin{bmatrix} x_{k-1} \\ w_{k-1} \end{bmatrix}; \begin{bmatrix} \hat{x}_{k-1} \\ 0 \end{bmatrix}, \begin{bmatrix} \tilde{P}_{k-1} & 0 \\ 0 & \tilde{Q}_{k-1} \end{bmatrix}, \tilde{\eta}_{k-1} \right) \tag{46}$$

as an approximate Student-*t* distribution. Since the degrees of freedom have changed, $\tilde{P}_{k-1}$ and $\tilde{Q}_{k-1}$ replace the matrices P_{k-1} and Q_{k-1}, and the mean remains unchanged. The method for finding $\tilde{P}_{k-1}$ and $\tilde{Q}_{k-1}$ called moment matching have already been given in Sec. 3.1.2. Therefore, we can now

apply a similar time update to (28)-(29). For measurement updates, a similar approximation is

$$p(x_k, v_k | \mathcal{Z}_{k-1})$$
$$\approx \mathrm{St}\left(\begin{bmatrix} x_k \\ v_k \end{bmatrix}; \begin{bmatrix} \hat{x}_{k|k-1} \\ 0 \end{bmatrix}, \begin{bmatrix} \tilde{P}_{k|k-1} & 0 \\ 0 & \tilde{R}_k \end{bmatrix}, \bar{\eta}_{k-1} \right) \tag{47}$$

The two density approximations (46) and (47) extend the exact filter in Section 3.1 and provide a convenient closed form solution for time update and measurement update. The approximate density is still the *t* density, but it may change the DOF.

Assuming that the measurement of each sensor is independent of each other, for each local sensor in the distributed sensor network, the following approximate consensus Student-*t* recursion process can be obtained:

(1) Prediction of local filter: The prediction update depends on the previous posterior parameters $\hat{x}_{k-1}$, P_{k-1} and η_{k-1}. First, the following approximation is required:

$$P_{k-1}^i, Q_{k-1}^i, \eta_{k-1}^i, \gamma_{k-1}^i \to \tilde{P}_{k-1}^i, \tilde{Q}_{k-1}^i, \tilde{\eta}_{k-1}^i \tag{48}$$

Then we have

$$\hat{x}_{k|k-1}^i = F_k \hat{x}_{k-1}^i \tag{49}$$

$$P_{k|k-1}^i = F_k \tilde{P}_{k-1}^i F_k^T + \tilde{Q}_{k-1} \tag{50}$$

(2) Update of local filter: Do the following approximation

$$P_{k|k-1}^i, R_k^i, \tilde{\eta}_{k-1}^i, \delta_k^i \to \tilde{P}_{k|k-1}^i, \tilde{R}_k^i, \bar{\eta}_{k-1}^i \tag{51}$$

Then we have

$$\tilde{z}_k^i = z_k^i - H_k^i \hat{x}_{k|k-1}^i \tag{52}$$

$$S_k^i = H_k^i \tilde{P}_{k|k-1}^i (H_k^i)^T + \tilde{R}_k^i \tag{53}$$

$$K_k^i = \tilde{P}_{k|k-1}^i (H_k^i)^T / S_k^i \tag{54}$$

$$x_{k,0}^i = x_{k|k-1}^i + K_k^i \tilde{z}_k^i \tag{55}$$

$$P_{k,0}^i = \frac{\bar{\eta}_{k-1}^i + \Delta_{z,k}^i}{\bar{\eta}_{k-1}^i + d_z^i} [\tilde{P}_{k|k-1}^i - K_k^i S_k^i (K_k^i)^T] \tag{56}$$

$$\Delta_{z,k}^i = (\tilde{z}_k^i)^T (S_k^i)^{-1} \tilde{z}_k^i \tag{57}$$

$$\eta_k^i = \bar{\eta}_{k-1}^i + d_z^i \tag{58}$$

(3) Consensus on the information matrix and information vector: Approximate the initial information matrices and information vectors by

$$\bar{P}_{k,0}^i = \frac{\eta_k^i}{\eta_k^i - 2} P_{k,0}^i \tag{59}$$

$$\Omega_{k,0}^i = (\bar{P}_{k,0}^i)^{-1} \tag{60}$$

$$q_{k,0}^i = \Omega_{k,0}^i \hat{x}_{k,0}^i \tag{61}$$

For a *L*-step consensus iteration, the consensus on posterior information is carried out in the form

$$\Omega_{k,\ell}^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} \Omega_{k,\ell-1}^j, \quad \ell = 1, \ldots, L \tag{62}$$

$$q_{k,\ell}^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} q_{k,\ell-1}^j, \quad \ell = 1, \ldots, L \tag{63}$$

where $\ell = 1, 2, \cdots, L$ is the consensus step, $\pi^{i,j}$ is the consensus weight. A convex combination is adopted by supposing $\pi^{i,j} \geq 0$ and $\sum_{j \in \mathcal{N}^i} \pi^{i,j} = 1$.

(4) Recover the state and the scale matrix:

$$\hat{x}_k^i = (\Omega_{k,L}^i)^{-1} q_{k,L}^i \tag{64}$$

$$P_k^i = \frac{\eta_k^i - 2}{\eta_k^i} (\Omega_{k,L}^i)^{-1} \tag{65}$$

C. DISCUSSION

Compared with the standard CI method, the growth of the computational complexity for the proposed method comes from calculating the $\Delta_{z,k}^i$ and η_k^i for exact method. The approximated method needs more computational cost in approximate procedures such as (48) and (51). Therefore, the order of computational complexity of the proposed algorithm is the same as that of CI method. In addition, no local parameters such as $\Delta_{z,k}^i$ and η_k^i are communicated among nodes, and only global parameters are communicated among neighboring nodes, the communication cost of the proposed method is the same as the standard CI method.

The Gaussian noise is one of the most common distribution in nature. The central limit theorem states that under appropriate conditions, the mean value of a large number of independent random variables converges to normal distribution according to the distribution after proper standardization. Under normal circumstances, the noise is generally Gaussian distribution, with occasional outliers which make the whole distribution of the noise have heavy-tailed feature. It can be seen from [36] that the abnormal values can also be expressed by Laplace distribution and Cauchy distribution. However, the Student-*t* distribution provides a heavy-tailed alternative to the Gaussian distribution while the shape of Student-*t* distribution is more similar to Gaussian distribution (see Fig. 1). When DOF tends to 1, the Student-*t* distribution becomes Cauchy distribution. Besides, the Student-*t* distribution leads to closed-form solution of the proposed filters. In addition, the Student-*t* based filter can deal with other heavy-tailed noises such as Laplace noise [22]. Therefore, we choose the Student-*t* distribution here to model the heavy-tailed process and measurement noises.
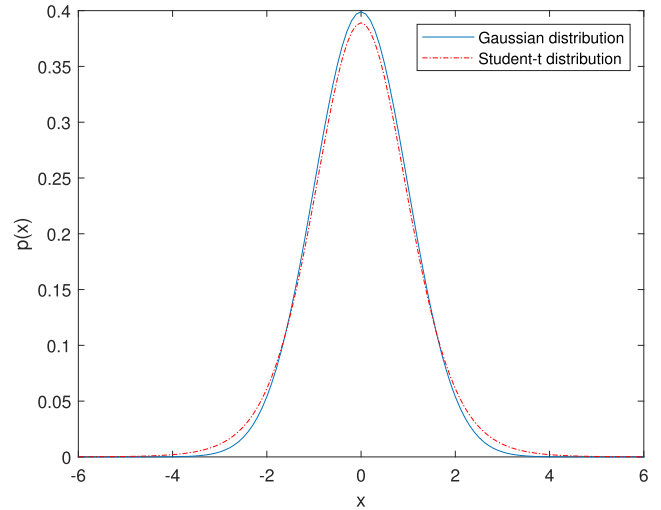
IV. SIMULATIONS

Here we consider a tracking problem in two-dimensional plane. The target dynamic includes the state $x = [p_x, \dot{p}_x, p_y, \dot{p}_y]^T$, and the model is confirmed by

$$F_k = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad Q_k = G_k \Delta G_k^T \tag{66}$$

where $\Delta = \text{diag}([w_x^2, w_y^2])$, sample time $T = 1$s, $w_x^2 = w_y^2 = 0.1$ and

$$G_k = \begin{bmatrix} T^2/2 & T & 0 & 0 \\ 0 & 0 & T^2/2 & T \end{bmatrix}^T \tag{67}$$

FIGURE 1. Gaussian distribution (with $m = 0$ and $P = 1$) and Student distribution (with $m = 0$, $P = 1$ and $v = 10$).

The true initial state of target is

$$x_0 = [2600\text{m}, 20\text{m/s}, 3800\text{m}, 10\text{m/s}]^T$$

Initial states for filters are chosen randomly from $N(x_0, P_0)$ in each simulations turn, where

$$P_0 = \text{diag}([50^2\text{m}^2, 5^2\text{m}^2/\text{s}^2, 50^2\text{m}^2, 5^2\text{m}^2/\text{s}^2])$$

There are 20 sensor nodes in the sensor network (the topology is shown in Fig. 2). The measurement model is given by

$$H_k^i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{68}$$

The heavy-tailed measurement noise is generated by a mixture of Gaussian with a nominal noise variance R and outliers with noise variance $100R$. Suppose we have a nominal measurement noise variance $R = \text{diag}([(10\text{m})^2, (10\text{m})^2])$,
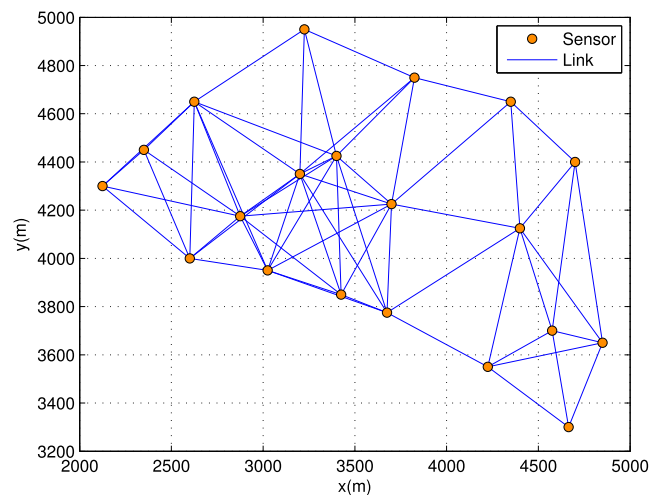
FIGURE 2. The sensor network used in simulations.

then the heavy-tailed measurement noise of sensor node i is generated according to

$$v_k^i \sim \begin{cases} \mathrm{N}(0, R), & \text{with probability } 1 - p_o \\ \mathrm{N}(0, 100R), & \text{with probability } p_o \end{cases} \quad (69)$$

where p_o is the probability of the measurement outliers. The heavy-tailed process noise is generated according to

$$w_k \sim \begin{cases} \mathrm{N}(0, Q), & \text{with probability } 1 - p_o \\ \mathrm{N}(0, 100Q), & \text{with probability } p_o \end{cases} \quad (70)$$

which is wildly used to evaluate the performance of Student-*t* based filters.

This paper mainly compares the following three methods: (1) Distributed consensus Kalman filter (DCKF) in [10]; (2) exact Distributed consensus Student-*t* filter, referred to as DCSTF-E; (3) approximate distributed consensus Student-*t* filter, referred to as DCSTF-A. The simulation results were obtained through 100 Monte Carlo simulations, and the root mean square error (RMSE) of position and velocity was used to evaluate the simulation results.

The consensus step is $L = 3$ and The consensus weights of sensor nodes are set to $\pi^{i,j} = 1/|\mathcal{N}^i|$ if $j \in \mathcal{N}^i$ and $\pi^{i,j} = 0$ if $j \notin \mathcal{N}^i$. The initial DOF of DCSTF-E are set to $\nu = 3$. For DCSTF-A, the DOF is set to $\gamma_k = \delta_k = \eta_0 = 20$.

When no outliers exist ($p_o = 0$), both process noise and measurement noise are Gaussian. Fig. 3 and Fig. 4 show the simulation results of the three methods without outliers. It can be seen from the figures that the RMSE of the DCKF and DCSTF-E methods are relatively close, and the RMSE of the DCSTF-A is slightly higher than the previous two methods. This means that in the absence of outliers, the performance of the DCKF and DCSTF-E methods is relatively close, while the performance of the DCSTF-A method is slightly lower.
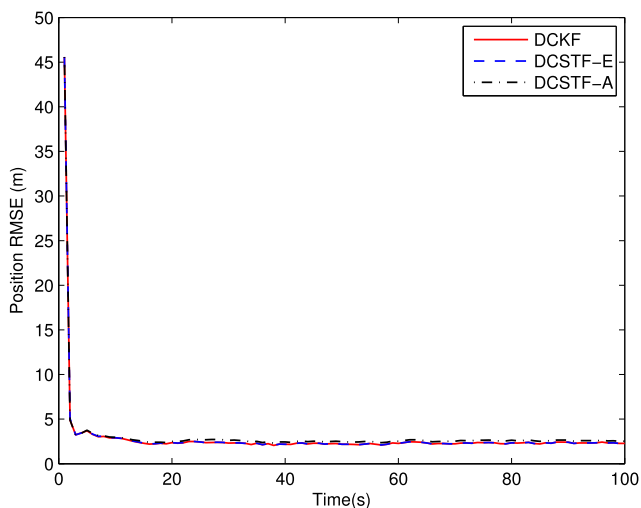

FIGURE 3. The Position RMSE for different methods when $p_o = 0$.

Fig. 5 and Fig. 6 show the simulation results of the three methods when the probability of outliers is $p_o = 0.2$. It can be seen from the figures that the RMSE of the DCSTF-E
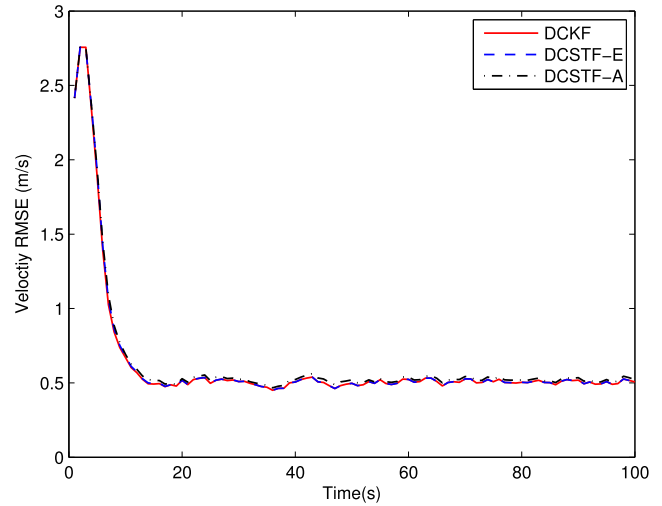

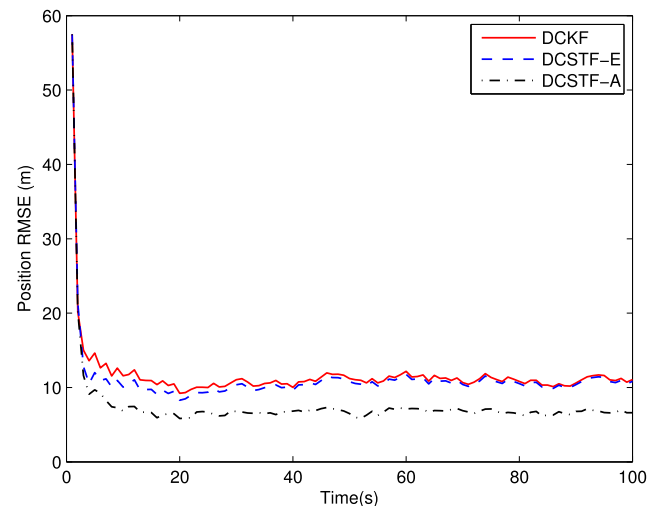**FIGURE 4.** The velocity RMSE for different methods when $p_o = 0$.

FIGURE 5. The Position RMSE for different methods when $p_o = 0.2$.

method at the initial stage of the simulation is lower than that of the DCKF method, and the root mean square error of the two is closer as the simulation proceeds. This is because the DOF parameter in the DCSTF-E method increases with time and eventually converges to the DCKF method. This shows that the previous theoretical analysis and simulation results are consistent. The RMSE of the DCSTF-A method is much lower than the first two methods. This shows that in the presence of outliers, the performance of the DCSTF-A method is significantly better than the DCKF and DCSTF-E methods.

Tab. 1 and Tab. 2 give the simulation results of the three methods under different outlier probabilities. It can be seen from the tables that as the outlier probability increases, the RMSE of all methods increases accordingly. Under different outlier probabilities, the RMSE of the DCSTF-E method is slightly lower than that of the DCKF method, and the RMSE of the DCSTF-A method is significantly lower than
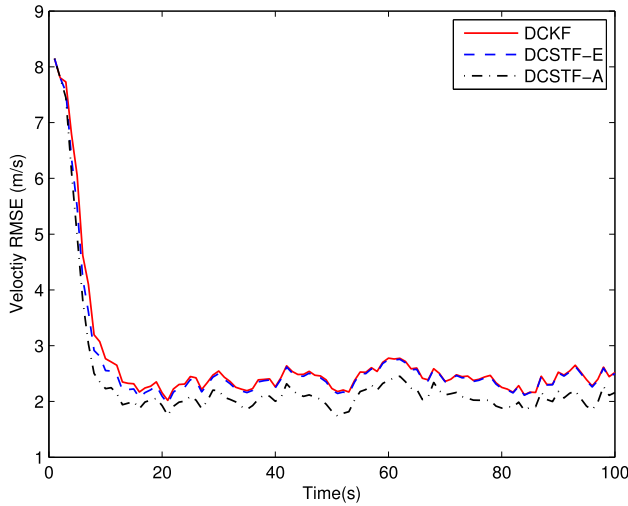
FIGURE 6. The velocity RMSE for different methods when $p_o = 0.2$.

TABLE 1. Position RMSE (m) with different outlier probabilities.

p_0	DCKF	DCSTF-E	DCSTF-A
0.1	8.3454	7.8950	5.9253
0.2	11.2848	10.6898	7.2796
0.3	14.2878	13.6191	9.3759
0.4	15.7395	15.0775	11.1987

TABLE 2. Velocity RMSE (m/s) with different outlier probabilities.

p_0	DCKF	DCSTF-E	DCSTF-A
0.1	1.7250	1.6759	1.5299
0.2	2.5601	2.5010	2.2131
0.3	3.2365	3.1810	2.7578
0.4	3.8304	3.7705	3.3223

TABLE 3. Position RMSE (m) with different consensus steps.

L	DCKF	DCSTF-E	DCSTF-A
1	13.7407	13.0275	9.0758
2	12.4748	11.8217	8.0229
3	11.2848	10.6898	7.3796
4	11.2058	10.6299	7.2041
5	10.8167	10.2834	7.0656

the first two methods. This shows that the proposed distributed consensus Student-*t* filtering algorithms perform better than the distributed consensus Kalman filtering algorithm in the presence of outliers.

Tab. 3 and Tab. 4 give the simulation results of the three methods under different consensus iteration steps. It can be seen from the tables that with the increase of the number of consensus iterations, the RMSE of the three methods decreases accordingly. Similar to the previous method, the RMSE of the DCSTF-E method is slightly lower than that of the DCKF method, and the RMSE of the DCSTF-A method is significantly lower than the previous two methods. These further verify the effectiveness of the proposed algorithms.

TABLE 4. Velocity RMSE (m/s) with different consensus steps.

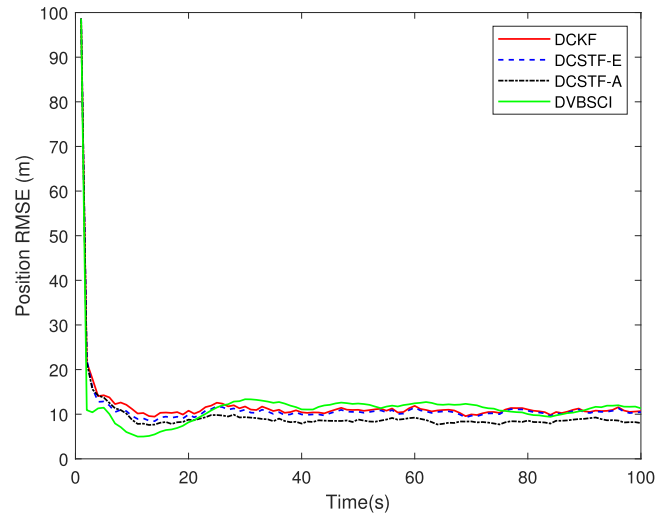| L	DCKF	DCSTF-E	DCSTF-A
1	2.6560	2.5896	2.2743
2	2.6136	2.5507	2.2340
3	2.5601	2.5010	2.2131
4	2.4988	2.4418	2.1490
5	2.4469	2.3948	2.1160

FIGURE 7. The Position RMSE for DVBSCI and other methods when $p_o = 0.2$.
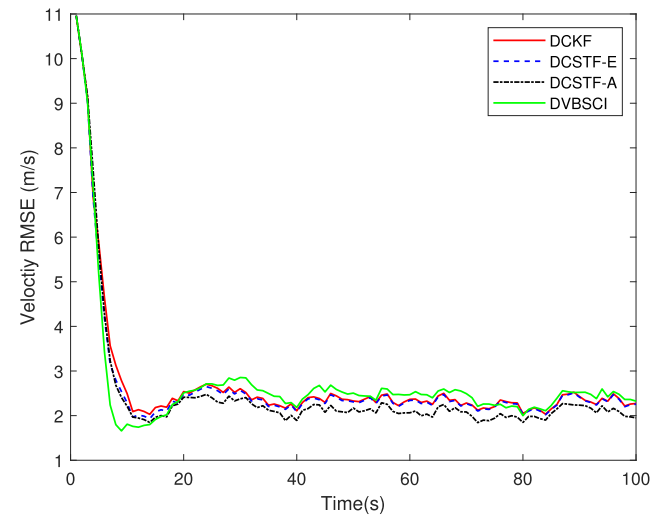

FIGURE 8. The velocity RMSE for DVBSCI and other methods when $p_o = 0.2$.

In order to further illustrate the effectiveness of the proposed algorithms, we the compared proposed algorithms with the distributed variational Bayesian Student-*t* CI (DVBSCI) filter presented in [25]. Fig. 7 and Fig. 8 show the simulation results. It can be seen from the figures that the RMSE of the DVBSCI method gives the worst performance. This is because DVBSCI algorithm only models the outliers of measurement noise and does not consider the outliers of process noise. This further proves the effectiveness of the

proposed algorithm considering outliers of both process noise and measurement noise.

V. CONCLUSION

In this chapter, a kind of consistent Student-t filter is proposed, which is used for both process noise and measurement noise. Firstly, the system model based on Student-t distribution is established. Based on the Student-t distribution, an accurate distributed consistent Student-t filter is proposed. Then, an approximate distributed consistent Student-t filter is proposed for the determination of the strong constraints of the filter and its convergence to the standard Kalman filter over time. The simulation results show that the proposed algorithm can achieve stable and accurate state estimation when both process noise and observation noise are heavy tailed noise. The proposed algorithm can be extended to other consensus algorithms and nonlinear situations in the future.

REFERENCES

[1] D. Ding, Q.-L. Han, Z. Wang, and X. Ge, "A survey on model-based distributed control and filtering for industrial cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2483–2499, May 2019.

[2] S. He, H.-S. Shin, S. Xu, and A. Tsourdos, "Distributed estimation over a low-cost sensor network: A review of state-of-the-art," *Inf. Fusion*, vol. 54, pp. 21–43, Feb. 2020.

[3] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. 44th IEEE Conf. Decis. Control*, Dec. 2005, pp. 8179–8184.

[4] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, Nov. 2010.

[5] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2069–2084, Sep. 2010.

[6] R. Olfati-Saber and J. S. Shamma, "Consensus filters for sensor networks and distributed sensor fusion," in *Proc. 44th IEEE Conf. Decis. Control*, Dec. 2005, pp. 6698–6703.

[7] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. Decis. Control*, Dec. 2007, pp. 5492–5498.

[8] G. Battistelli, L. Chisci, S. Morrocchi, and F. Papi, "An information-theoretic approach to distributed state estimation," *IFAC Proc. Volumes*, vol. 44, no. 1, pp. 12477–12482, 2011.

[9] G. Battistelli, L. Chisci, and C. Fantacci, "Parallel consensus on likelihoods and priors for networked nonlinear filtering," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 787–791, Jul. 2014.

[10] G. Battistelli and L. Chisci, "Kullback–Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707–718, Mar. 2014.

[11] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano, "Consensus-based linear and nonlinear filtering," *IEEE Trans. Autom. Control*, vol. 60, no. 5, pp. 1410–1415, May 2015.

[12] K. Shen, Z. Jing, and P. Dong, "A consensus nonlinear filter with measurement uncertainty in distributed sensor networks," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1631–1635, Nov. 2017.

[13] M. A. Jacobs and D. DeLaurentis, "Distributed Kalman filter with a Gaussian process for machine learning," in *Proc. IEEE Aerosp. Conf.*, Mar. 2018, pp. 1–12.

[14] M. Zorzi, "Distributed Kalman filtering under model uncertainty," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 2, pp. 990–1001, Jun. 2020.

[15] J. Zhou, G. Gu, and X. Chen, "Distributed Kalman filtering over wireless sensor networks in the presence of data packet drops," *IEEE Trans. Autom. Control*, vol. 64, no. 4, pp. 1603–1610, Apr. 2019.

[16] S. P. Talebi and S. Werner, "Distributed Kalman filtering and control through embedded average consensus information fusion," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4396–4403, Oct. 2019.

[17] J. Wang, P. Dong, Z. Jing, and J. Cheng, "Consensus variable structure multiple model filtering for distributed maneuvering tracking," *Signal Process.*, vol. 162, pp. 234–241, Sep. 2019.

[18] J.-A. Ting, E. Theodorou, and S. Schaal, "A Kalman filter for robust outlier detection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2007, pp. 1514–1519.

[19] M. Roth, E. Ozkan, and F. Gustafsson, "A Student's t filter for heavy tailed process and measurement noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 5770–5774.

[20] J. Christmas and R. Everson, "Robust autoregression: Student-t innovations using variational bayes," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 48–57, Jan. 2011.

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

[22] H. Zhu, H. Leung, and Z. He, "A variational Bayesian approach to robust sensor fusion based on Student-t distribution," *Inf. Sci.*, vol. 221, pp. 201–214, Feb. 2013.

[23] P. Dong, Z. Jing, H. Leung, K. Shen, and M. Li, "The labeled multi-Bernoulli filter for multitarget tracking with glint noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2253–2268, Oct. 2019.

[24] Z. Zhang, P. Dong, J. Wang, and Y. Sun, "Improving S-MSCKF with variational Bayesian adaptive nonlinear filter," *IEEE Sensors J.*, pp. 1–12, 2020.

[25] P. Dong, Z. Jing, K. Shen, and M. Li, "A distributed consensus filter for sensor networks with heavy-tailed measurement noise," *Sci. China Inf. Sci.*, vol. 61, no. 11, Nov. 2018, Art. no. 119201.

[26] P. Dong, Z. Jing, H. Leung, and K. Shen, "Variational Bayesian adaptive cubature information filter based on wishart distribution," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 6051–6057, Nov. 2017.

[27] J. Li, F. Deng, and J. Chen, "A fast distributed variational Bayesian filtering for multisensor LTV system with non-Gaussian noise," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2431–2443, Jul. 2019.

[28] P. Dong, Z. Jing, H. Leung, K. Shen, and M. Li, "Robust consensus nonlinear information filter for distributed sensor networks with measurement outliers," *IEEE Trans. Cybern.*, vol. 49, no. 10, pp. 3731–3743, Oct. 2019.

[29] N. J. Gordon and A. F. M. Smith, "Approximate non-Gaussian Bayesian estimation and modal consistency," *J. Roy. Stat. Soc. B, Methodol.*, vol. 55, no. 4, pp. 913–918, Sep. 1993.

[30] I. Bilik and J. Tabrikian, "Maneuvering target tracking in the presence of glint using the nonlinear Gaussian mixture Kalman filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 246–262, Jan. 2010.

[31] Y. Chen, F. Liu, S. Mei, and J. Ma, "A robust WLAV state estimation using optimal transformations," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2190–2191, Jul. 2015.

[32] Y. Chen, J. Ma, P. Zhang, F. Liu, and S. Mei, "Robust state estimator based on maximum exponential absolute value," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1537–1544, Jul. 2017.

[33] O. Straka and J. Dunik, "Stochastic integration student's-t filter," in *Proc. 20th Int. Conf. Inf. Fusion (Fusion)*, Jul. 2017, pp. 1–8.

[34] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers, "A novel robust student's T-based Kalman filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1545–1554, Jun. 2017.

[35] P. Dong, Z. Jing, H. Leung, K. Shen, and J. Wang, "Student-T mixture labeled multi-Bernoulli filter for multi-target tracking with heavy-tailed noise," *Signal Process.*, vol. 152, pp. 331–339, Nov. 2018.

[36] Y. Li, J. Li, J. Qi, and L. Chen, "Robust cubature Kalman filter for dynamic state estimation of synchronous machines under unknown measurement noise statistics," *IEEE Access*, vol. 7, p. 29139–29148, 2019.
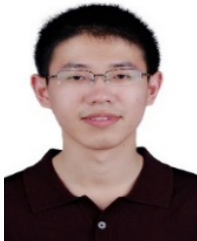
JINRAN WANG received the B.S. and master's degrees from the Beijing Institute of Technology, Beijing, China, in 2013 and 2016, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2019.

He is currently an Engineer with the Beijing Institute of Electronic System Engineering. His research interests include information fusion and nonlinear filtering.

PENG DONG was born in Pingyi, Linyi, China, in 1985. He received the B.S. and Ph.D. degrees from Northwestern Polytechnical University, Xi'an, in 2008 and 2013, respectively.

From 2013 to 2015, he was a Postdoctoral Fellow with Shanghai Jiao Tong University. He is currently an Associate Professor with the School of Aeronautics and Astronautics, Shanghai Jiao Tong University. His research interests include information fusion, SLAM, target tracking, and nonlinear filtering.

XUN SONG was born in Nanyang, Henan, China, in 1984. He received the B.S. degree from the Beijing Institute of Graphic Communication, Beijing, in 2006, and the Ph.D. degree from Peking University, Beijing, in 2013.

From 2011 to 2012, he was a Visiting Scholar with North Carolina State University. He is currently a Senior Engineer with the Beijing Institute of Electronic System Engineering. His research interests include information fusion and swarm intelligence.

KAI SHEN was born in Fujian, China. He received the B.S. and M.S. degrees from Northwestern Polytechnical University, Xi'an, China, in 2011 and 2014, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2019.

He is currently an Assistant Professor with the School of Electrical Engineering, Southwest Jiaotong University. His research interests include target tracking and information fusion.

XIAODONG WANG was born in Jinzhong, Shanxi, China, in 1982. He received the B.S. degree from Northwestern Polytechnical University, Xi'an, China, in 2004, and the M.S. degree from the Beijing Institute of Electronic System Engineering, Beijing, China, in 2007.

He is currently a Senior Engineer with the Beijing Institute of Electronic System Engineering. His research interests include information fusion and swarm intelligence.

● ● ●