

Article

Distributed Consensus Multi-Distribution Filter for Heavy-Tailed Noise

Guan-Nan Chang^{1,2}, Wen-Xing Fu¹, Tao Cui³, Ling-Yun Song³  and Peng Dong^{3,*} 

¹ Unmanned System Research Institute, Northwestern Polytechnical University, Xi'an 710072, China; gnchang2019@163.com (G.-N.C.); wenxingfu@nwpu.edu.cn (W.-X.F.)

² Xi'an Modern Control Technology Research Institute, Xi'an 710065, China

³ School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, China; cui tao@sjtu.edu.cn (T.C.); songlingyun@sjtu.edu.cn (L.-Y.S.)

* Correspondence: dongpengkty@sjtu.edu.cn

Abstract: Distributed state estimation is one of the critical technologies in the field of target tracking, where the process noise and measurement noise may have a heavy-tailed distribution. Traditionally, heavy-tailed distributions like the student-t distribution are employed, but our observation reveals that Gaussian noise predominates in many instances, with occasional outliers. This sporadic reliance on heavy-tailed distributions can degrade performances or necessitate frequent parameter adjustments. To overcome this, we introduce a novel distributed consensus multi-distribution state estimation method that combines Gaussian and student-t filters. Our approach establishes a system model using both Gaussian and student-t distributions. We derive a multi-distribution filter for a single sensor, assigning probabilities to Gaussian and student-t noise models. Parallel estimation under both distributions, utilizing Gaussian and student-t filters, allows us to calculate the likelihood of each distribution. The fusion of these results yields a mixed-state estimation and corresponding error matrix. Recognizing the increasing degrees of freedom in the student-t distribution over time, we provide an effective approximation. An information consensus strategy for multi-distribution filters is introduced, achieving global estimation through consensus on fused local filter results via interaction with neighboring nodes. This methodology is extended to the distributed case, and the recursive process of the distributed multi-distribution consensus state estimation method is presented. Simulation results demonstrate that the estimation accuracy of the proposed algorithm improved by at least 20% compared to that of the traditional algorithm in scenarios involving both Gaussian and heavy-tailed distributions.

Keywords: multi-distribution; consensus filters; distributed state estimation; sensor networks; student-t distribution; heavy-tailed distribution



Citation: Chang, G.-N.; Fu, W.-X.; Cui, T.; Song, L.-Y.; Dong, P. Distributed Consensus Multi-Distribution Filter for Heavy-Tailed Noise. *J. Sens. Actuator Netw.* **2024**, *13*, 38.

<https://doi.org/10.3390/jsan13040038>

<https://doi.org/10.3390/jsan13040038>

jsan13040038

Academic Editor: Lei Shu

Received: 8 April 2024

Revised: 26 June 2024

Accepted: 26 June 2024

Published: 28 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Background

Sensor networks have garnered significant attention in recent years in distributed applications, such as environmental monitoring, power grid systems and traffic management [1–3]. Distributed target tracking or state estimation is one of the classical applications, and the core of it is the use of local communication and measurement through information exchange to obtain the state estimation of the entire system. However, some interference may lead to heavy-tailed measurement noise, as shown in Figure 1, which could degrade the estimation performance. Thus, how to improve the estimation performance in distributed scenarios with heavy-tailed noise is a troublesome problem.

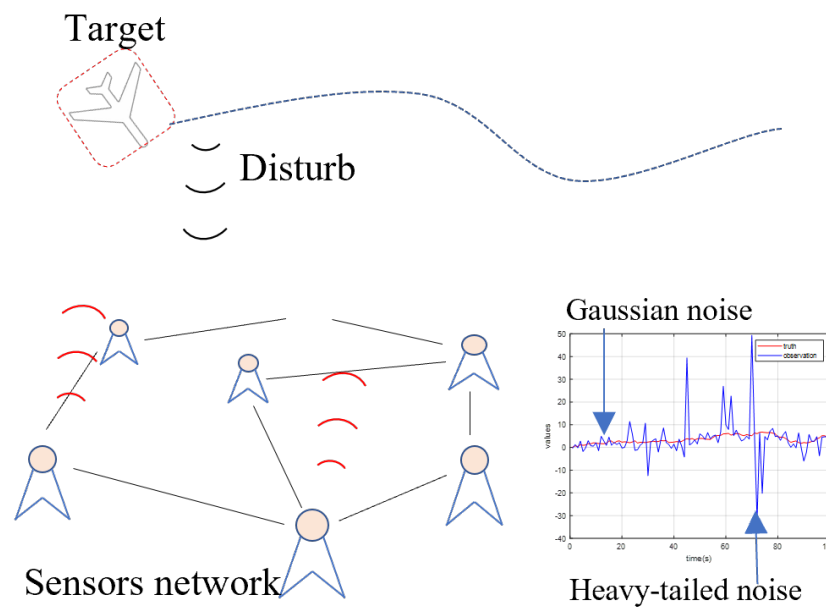


Figure 1. Research background and problem.

1.2. Related Work

There are three main types of distributed state estimation methods, including consensus methods, gossip methods and diffusion methods. For gossip-based method [4,5], each sensor selects one or several connected neighbors randomly to send information to. Although they have a low communication cost, gossip-based methods show slow convergence. The diffusion methods use a one-step convex combination of the received data [6,7], so they cannot converge to a consensus result. In contrast, the consensus method can ensure global convergence in many cases [8,9]; thus, it is the most widely used. Combining this method with Kalman filtering is the most direct method, and there are four main types: the consensus on estimation (CE) method, consensus on measurement (CM) method, consensus on information (CI) method and hybrid CM-CI (HCMCI) method [10,11]. Since the CE method does not consider the error covariance matrix, the estimation results are conservative [12]. In order to solve this problem, the CM method is proposed. However, the bounded estimation error of the CM method can only be achieved when consensus steps are large enough [8]. The CI method conforms the information vector and the information matrix, which can ensure the bounding of the estimation error of any number of consensus steps, while the new information is inevitably under-weighted [13–15]. The HCMCI method can handle both a priori and likelihood consensus, combining the advantages of the CM and CI methods. In addition, there are consensus algorithms [16] and consensus algorithms based on the Luenberger observer [17].

However, the above systems assume that the system noise obeys a Gaussian distribution, which is easy to be violated in practice. Due to unmodeled anomalies, sudden disturbances from the environment, unreliable sensors, target maneuvers, system failures or attacks on the system, the system will suffer from outliers [18]. In such scenarios, relying solely on the Gaussian noise assumption for state estimation leads to severe performance degradation and potential system divergence [19]. Outliers in distributions typically exhibit heavy-tailed characteristics, prompting the utilization of heavy-tailed distributions to accurately characterize system states and measurements [20]. The student-t distribution serves as an effective model for simulating heavy-tailed distributions, especially in handling non-Gaussian outliers, rendering it a widely adopted solution in this context [21]. In [22], the sensor measurement noise was modeled as a multivariable student-t distribution using the CI consensus strategy, where the state and noise parameters are estimated at the same

time. Combined with the variational Bayesian (VB) method, the joint posterior distribution is processed.

However, the above method can only deal with the situation where the measurement noise follows a heavy-tailed distribution. In order to deal with the situation that both process noise and measurement noise follow a heavy-tailed distribution, ref. [23] established both process and measurement noises as student-t distributions and proposed a distributed state estimation method based on a student-t filter and CI consensus strategy. In practical real-world settings, noise predominantly adheres to a Gaussian distribution, with outliers occurring sporadically. Given this scenario, an intuitive approach involves modeling the system where a mixture of Gaussian and student-t distributions encapsulates the dynamic nature of the noise. This method capitalizes on the strengths inherent in both Gaussian and student-t distributions, aiming to attain precise and consistent state estimations. Grounded in this conceptual framework, a resilient filtering algorithm, utilizing a blend of Gaussian and student-t distributions, is introduced for single-sensor state estimation [24]. Nevertheless, the iterative calculations of the parameters pose a challenge, particularly in the context of distributed systems characterized by lightweight architectures and limited communication bandwidth.

We can find that many studies have focused on the robustness. However, in practical scenarios, heavy-tailed noise is a low-probability event, and the model noise can still present as a Gaussian distribution in most cases. Although the mentioned robust filter can provide a robust performance in the presence of outliers, it will lose estimation accuracy when the noise is normal. Therefore, how to balance the robustness and estimation accuracy as well as obtain consensus results are significant problems.

Motivated by these, this paper proposes a distributed consensus state estimation method based on the Gaussian distribution and the student-t distribution. A comparison of related studies is shown in Table 1.

Table 1. Comparison with related studies.

Reference	Consensue	Gaussian Distribution	Heavy-Tailed Distribution
[10–12,16,17]	✓	✓	
[19–21]			✓
[22,23]	✓		✓
[24]		✓	✓
the proposed algorithm	✓	✓	✓

The main contribution of this paper can be highlighted as follows:

- (1) The derivation of a method for local sensors that models process and measurement noises in both Gaussian and student-t distributions. Each distribution is assigned a probability, enabling parallel state estimation using Gaussian and student-t filters. The fusion of the results from these distributions provides a mixed-state estimation and corresponding error matrix.
- (2) Addressing the challenge of the increasing degrees of freedom in the student-t distribution over time by providing an effective approximate solution, ensuring stability and accuracy in the estimation process.
- (3) The introduction of an information consensus strategy for multi-distribution filters, enabling global estimation by achieving consensus on fused local filter results through interaction with neighboring nodes. The results are extended to the distributed case, and the recursive process presented further validates its efficacy, as supported by simulation results showcasing its performance in scenarios involving both Gaussian and heavy-tailed distributions.

The organization of this paper is shown in Figure 2. Section 2 presents the problem, Section 3 is the proposed distributed multiple-consensus-state estimation algorithm, the simulation results are presented in Section 4, and the conclusions are presented in Section 5.

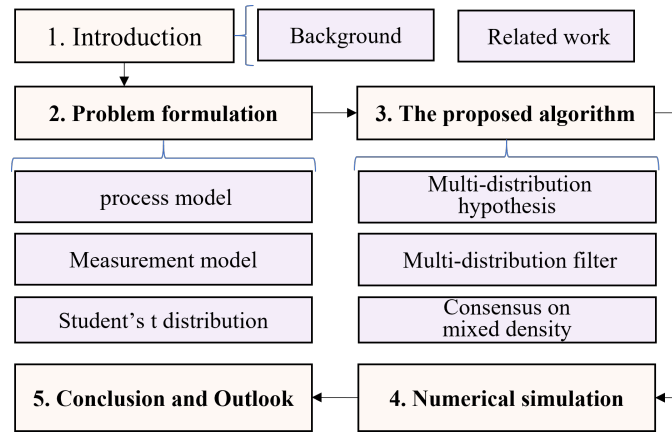


Figure 2. The organization of this paper.

2. Preliminaries and Problem Formulation

We consider a sensor network with node \mathcal{N} and edge \mathcal{A} . Then, the network can be represented as a two-tuple $(\mathcal{N}, \mathcal{A})$. The set of neighbor nodes (including node i itself) of node $i \in \mathcal{N}$ is \mathcal{N}^i . The process equation of the system is assumed to be discrete-time linear:

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

where x_k means the n_x -dimension state vector at time k , F_k means the state transition matrix, and w_k means the process noise at time k .

The measurement equation of each node $i \in \mathcal{N}$ is also linear:

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

where z_k^i means the n_z^i -dimension measurement vector of node $i \in \mathcal{N}$ at time k , H_k^i means the measurement matrix of node $i \in \mathcal{N}$, and v_k^i means the measurement noise of node $i \in \mathcal{N}$ at time k .

In general, the noises can be supposed to follow a Gaussian distribution. However, some unknown perturbations may lead to outliers characterized by heavy-tailed noise. In order to deal with the problem, many studies have used distributions that are insensitive to outliers to model the noise.

The student-t distribution is one of them. A random variable x obeys the student-t distribution if its probability density satisfies the following [20]:

$$\text{St}(x; m, P, \eta) = M \left(1 + \frac{1}{\eta} (x - m)^T P^{-1} (x - m)\right)^{-\frac{\eta+n_x}{2}} \tag{3}$$

where

$$M = \frac{\Gamma(\frac{\eta+n_x}{2})}{\Gamma(\frac{\eta}{2})} \frac{1}{(\eta\pi)^{\frac{n_x}{2}}} \frac{1}{\sqrt{\det(P)}} \tag{4}$$

m is the mean, η is the degrees of freedom (dof), P is a positive definite symmetric matrix, Γ denotes the Gamma function, and n_x is the dimension of x . The variables of the system model are summarized in Table 2.

Table 2. Variables of system model.

Variables	Meaning	Variables	Meaning
x	State	P	Covariance
z	Measurement	η	dof
m	Mean value	Γ	Gamma function

It should be noted that the variance in the student-t distribution is $\frac{\eta}{\eta-2}P$ for dof η greater than 2; otherwise, the variance is undefined. When $\eta = 1$, it converges to the Cauchy distribution, and while η tends to infinity, it becomes the Gaussian distribution. Compared with the Gaussian distribution, the student-t distribution shows heavy-tailed characteristics: when moving away from the mean, the density does not decrease as quickly as with the Gaussian distribution. Figure 3 presents logarithmic presentations of the Gaussian and student-t distribution. It can be seen from the figure that the value of the Gaussian distribution quickly drops below -16 , while the value of the student-t distribution does not drop so fast.

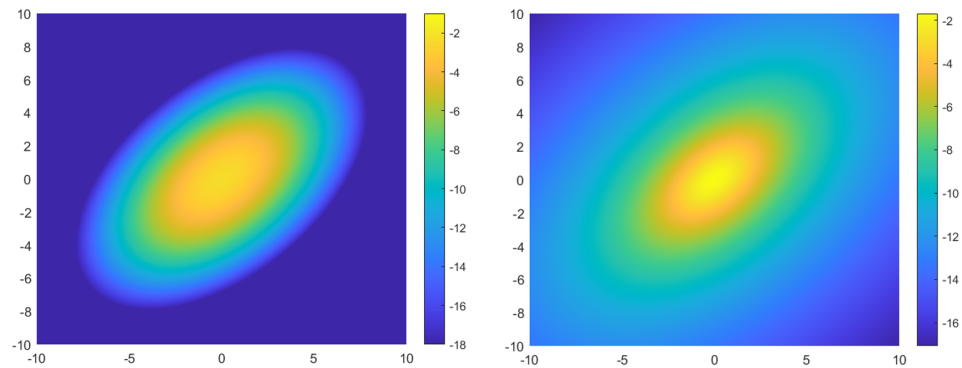


Figure 3. Logarithmic heat maps of Gaussian distribution (left) and student-t distribution (right).

Due to the long-tailed property, the student-t distribution is effective in dealing with heavy-tailed noise. However, it may not describe normal noise well. In practical scenarios, the Gaussian noise can adapt to most situations, while the heavy-tailed noise is a low-probability event. Thus, only student-t distribution molding can improve the robustness of the algorithm, but it will loss estimation accuracy.

The problem is how to balance and obtain the robustness and estimation accuracy of the consensus estimation results for sensor networks in the presence of possible heavy-tailed noise.

3. Main Results

The proposed multi-distribution filter, based on Gaussian and student-t distributions for a single local sensor operating without information interaction, is introduced, addressing inherent algorithm limitations. An approximate method is provided to mitigate these shortcomings. Subsequently, a CI-based consensus strategy is proposed for scenarios involving mixed Gaussian and student-t distributions. Finally, building upon the aforementioned consensus strategy and the single-sensor multi-distribution filter, an algorithm for distributed multi-distribution filtering is proposed.

3.1. The Multi-Distribution Filter Based on Gaussian Distribution and Student-t Distribution

In this subsection, we present a multi-distribution filter for a single sensor exposed to both a heavy-tailed process and measurement noises. The exclusive reliance on a student-t distribution filter often results in prolonged performance degradation or necessitates frequent parameter readjustments during normal system operations. Conversely, relying solely on Gaussian filters tends to cause divergence in scenarios where system outliers manifest. To harness the strengths of both filters, we present the following two hypotheses for a single sensor node (omitting the superscript representing the sensor node in this subsection):

H_0 : Suppose that the process and measurement noises obey the Gaussian distribution as follows:

$$p(w_k) = N(w_k; 0, Q_k^0) \tag{5}$$

$$p(v_k) = N(v_k; 0, R_k^0) \tag{6}$$

where $N(x; m, P)$ denotes that x obeys the Gaussian distribution with mean m and covariance P .

H_1 : Suppose that the process and measurement noises obey the student-t distribution as follows:

$$p(w_k) = \text{St}(w_k; 0, Q_k^1, \eta_k) \tag{7}$$

$$p(v_k) = \text{St}(v_k; 0, R_k^1, \eta_k) \tag{8}$$

where the probability of H_0 is μ_k^0 , the probability of H_1 is μ_k^1 , $\sum_{r \in \mathcal{H}} \mu_k^r = 1$, and $\mathcal{H} = \{0, 1\}$.

Now, we need to assign a filter to each distribution. For the hypothesis of the Gaussian distribution, since the system model is linear, the standard Kalman filter can be used. The steps are as follows: given the initial values \hat{x}_0^0 and \hat{P}_0^0 , when time step $k > 1$, the following recursive process is performed:

(1) Time update

$$\hat{x}_{k|k-1}^0 = F_k \hat{x}_{k-1}^0 \tag{9}$$

$$P_{k|k-1}^0 = F_k P_{k-1}^0 F_k^T + Q_{k-1}^0 \tag{10}$$

(2) Measurement update

$$\tilde{z}_k^0 = z_k^0 - H_k \hat{x}_{k|k-1}^0 \tag{11}$$

$$S_k^0 = H_k P_{k|k-1}^0 (H_k)^T + R_k^0 \tag{12}$$

$$K_k^0 = P_{k|k-1}^0 (H_k)^T / S_k^0 \tag{13}$$

$$\hat{x}_k^0 = \hat{x}_{k|k-1}^0 + K_k^0 \tilde{z}_k^0 \tag{14}$$

$$P_k^0 = P_{k|k-1}^0 - K_k^0 S_k^0 (K_k^0)^T \tag{15}$$

For the hypothesis of the student-t distribution, we use the student-t filter described as follows: given the initial values \hat{x}_0^1 , P_0^1 and η_0 , when time step $k > 1$, the following recursive process is performed:

(1) Time update

$$\hat{x}_{k|k-1}^1 = F_k \hat{x}_{k-1}^1 \tag{16}$$

$$P_{k|k-1}^1 = F_k P_{k-1}^1 F_k^T + Q_{k-1}^1 \tag{17}$$

(2) Measurement update

$$\tilde{z}_k^1 = z_k^1 - H_k \hat{x}_{k|k-1}^1 \tag{18}$$

$$S_k^1 = H_k P_{k|k-1}^1 (H_k)^T + R_k^1 \tag{19}$$

$$K_k^1 = P_{k|k-1}^1 (H_k)^T / S_k^1 \tag{20}$$

$$\hat{x}_k^1 = \hat{x}_{k|k-1}^1 + K_k^1 \tilde{z}_k^1 \tag{21}$$

$$P_k^1 = \frac{\eta_{k-1} + \Delta_{z,k}^1}{\tilde{\eta}_{k-1} + d} (P_{k|k-1}^1 - K_k^1 S_k^1 (K_k^1)^T) \tag{22}$$

$$\Delta_{z,k}^1 = (\tilde{z}_k^1)^T (S_k^1)^{-1} \tilde{z}_k^1 \tag{23}$$

$$\eta_k = \eta_{k-1} + n_z \tag{24}$$

The detailed derivation of the student-t filter can be seen in [25].

In assuming that the state posterior obeys the mixed distribution of Gaussian and student-t distributions, their probabilities are μ_k^0 and μ_k^1 , respectively. According to the full probability theorem,

$$\begin{aligned} p(x_k | Z^k) &= \sum_{r \in \mathcal{H}} \mu_k^r p(x_k | Z^k, H_r) \\ &= \mu_k^0 N(x_k; \hat{x}_k^0, P_k^0) + \mu_k^1 \text{St}(x_k; \hat{x}_k^1, P_k^1, \eta_k) \end{aligned} \tag{25}$$

where $Z^k = \{z_1, z_2, \dots, z_k\}$ represents the measurement set up to time k . In order to obtain the posterior probability density function (PDF), the required parameters are the probability μ_k^r corresponding to the two distributions, the state estimation \hat{x}_k^r and the matrix P_k^r . For the student-t distribution, the dof η_k is also required. In addition to the distribution probability μ_k^r , the others are obtained by the two parallel filters. Given the distribution probability μ_{k-1}^i of the previous time, the probability of assuming that H_r is correct can be given by

$$\begin{aligned} \mu_k^r &\triangleq p(H_r|Z^k) = p(H_r|z_k, Z^{k-1}) \\ &= \frac{p(z_k|Z^{k-1}, H_r)p(H_r|Z^{k-1})}{p(z_k|Z^{k-1})} \\ &= \frac{p(z_k|Z^{k-1}, H_r)p(H_r|Z^{k-1})}{\sum_{j \in \mathcal{H}} p(z_k|Z^{k-1}, H_j)p(H_j|Z^{k-1})} \\ &= \frac{p(z_k|Z^{k-1}, H_r)\mu_{k-1}^r}{\sum_{j \in \mathcal{H}} p(z_k|Z^{k-1}, H_j)\mu_{k-1}^j} \end{aligned} \tag{26}$$

where $p(z_k|Z^{k-1}, H_r)$ is the measurement likelihood of H_r at time k , that is, $\Lambda_k^r \triangleq p(z_k|Z^{k-1}, H_r) = p(\tilde{z}_k^r)$. For the Gaussian filter,

$$\Lambda_k^0 = \frac{1}{\sqrt{\det(2\pi S_k^0)}} e^{-\frac{1}{2}(\tilde{z}_k^0)^T (S_k^0)^{-1} \tilde{z}_k^0} \tag{27}$$

For the student-t distribution filter,

$$\Lambda_k^1 = \frac{\Gamma(\frac{\eta_{k-1} + n_z}{2})}{\Gamma(\frac{\eta_{k-1}}{2})} \frac{1}{(\eta_{k-1}\pi)^{\frac{n_z}{2}}} \frac{1}{\sqrt{\det(S_k^1)}} \left(1 + \frac{1}{\eta_{k-1}} (\tilde{z}_k^1)^T (S_k^1)^{-1} \tilde{z}_k^1\right)^{-\frac{\eta_{k-1} + n_z}{2}} \tag{28}$$

The mean of the mixed posterior distribution is

$$\begin{aligned} \hat{x}_k &= E[x_k|Z^k] = \int x_k [\mu_k^0 \mathbf{N}(x_k; \hat{x}_k^0, P_k^0) + \mu_k^1 \mathbf{St}(x_k; \hat{x}_k^1, P_k^1, \eta_k)] dx_k \\ &= \mu_k^0 \int x_k \mathbf{N}(x_k; \hat{x}_k^0, P_k^0) dx_k + \mu_k^1 \int x_k \mathbf{St}(x_k; \hat{x}_k^1, P_k^1, \eta_k) dx_k \\ &= \mu_k^0 E[x_k|Z^k, H_0] + \mu_k^1 E[x_k|Z^k, H_1] \\ &= \sum_{r \in \mathcal{H}} \mu_k^r \hat{x}_k^r. \end{aligned} \tag{29}$$

The covariance corresponding to the mixed posterior distribution is

$$\begin{aligned} P_k &= E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T | Z^k] \\ &= \int (x_k - \hat{x}_k)(x_k - \hat{x}_k)^T [\mu_k^0 \mathbf{N}(x_k; \hat{x}_k^0, P_k^0) + \mu_k^1 \mathbf{St}(x_k; \hat{x}_k^1, P_k^1, \eta_k)] dx_k \\ &= \mu_k^0 \int (x_k - \hat{x}_k)(x_k - \hat{x}_k)^T \mathbf{N}(x_k; \hat{x}_k^0, P_k^0) dx_k + \mu_k^1 \int (x_k - \hat{x}_k)(x_k - \hat{x}_k)^T \mathbf{St}(x_k; \hat{x}_k^1, P_k^1, \eta_k) dx_k \\ &= \sum_{r \in \mathcal{H}} \mu_k^r E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T | Z^k, H_r] \\ &= \sum_{r \in \mathcal{H}} \mu_k^r E[(x_k - \hat{x}_k^r + \hat{x}_k^r - \hat{x}_k)(x_k - \hat{x}_k^r + \hat{x}_k^r - \hat{x}_k)^T | Z^k, H_r] \\ &= \sum_{r \in \mathcal{H}} \mu_k^r E[(x_k - \hat{x}_k^r)(x_k - \hat{x}_k^r)^T | Z^k, H_r] + \sum_{r \in \mathcal{H}} \mu_k^r (\hat{x}_k^r - \hat{x}_k)(\hat{x}_k^r - \hat{x}_k)^T. \end{aligned} \tag{30}$$

For the Gaussian distribution, the variance is

$$P_k^0 = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T | Z^k, H_0] \tag{31}$$

and for the student-t distribution, the variance is

$$\bar{P}_k^1 = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T | Z^k, H_1] = \frac{\eta_k}{\eta_k - 2} P_k^1. \tag{32}$$

Therefore, the expression of the covariance of the mixed posterior distribution can be obtained by

$$P_k = \mu_k^0 P_k^0 + \mu_k^1 \bar{P}_k^1 + \sum_{r \in \mathcal{H}} \mu_k^r (\hat{x}_k^r - \hat{x}_k)(\hat{x}_k^r - \hat{x}_k)^T \tag{33}$$

Thus, we obtain a complete recursive step for the multi-distribution filter.

With each measurement update, the dof increase according to Equation (24). In turn, this also requires an increase in the dof of noise, making the problem more and more Gaussian. In fact, the algorithm will converge to the Kalman filter after several time steps. Therefore, it is necessary to find some approximate methods.

One of the simplest ways is to enforce $\eta_k = \eta$ in Equation (24), where η is a constant, so that the dof will not increase all the time. However, the actual posterior density is $p(x_k) = \text{St}(x_k; \hat{x}_k^1, P_k^1, \eta + n_z)$ instead of $\text{St}(x_k; \hat{x}_k^1, P_k^1, \eta)$. Notice that we have omitted the condition here. At this point, we need to find a posterior density $q(x_k) = \text{St}(x_k; \hat{x}_k^1, \bar{P}_k^1, \eta)$ to approximate $p(x_k)$. The qualitative characteristics should be retained. Therefore, the adjusted matrix parameters should be the scaled version of the original matrix. The general expression of this problem is how to find the probability density $q(x) = \text{St}(x; \hat{x}, cP, \tilde{\eta})$ to approximate the probability density $p(x) = \text{St}(x; \hat{x}, cP, \eta)$ so that the two are as close as possible. This density is controlled by the scaling factor $c > 0$. Therefore, our problem is how to find c so that the two probability densities are close in a certain sense. A suggested method is to use the moment matching method, that is, to make the variances of the $p(x)$ and $q(x)$ equal. The advantages of this method are that it is simple to apply and does not need any parameters to be adjusted. We can obtain the following conditions:

$$\frac{\eta}{\eta - 2} P = \frac{\tilde{\eta}}{\tilde{\eta} - 2} cP \tag{34}$$

where $\eta > 2$ and $\tilde{\eta} > 2$. Then, the scale factor c can be obtained:

$$c = \frac{\eta(\tilde{\eta} - 2)}{(\eta - 2)\tilde{\eta}} \tag{35}$$

The PDF of process and measurement noises can be approximated in the same way.

3.2. Consensus on Mixed Density

The basic idea of consensus is to calculate the aggregation of the whole network by iteratively computing the same type of region on each node in the network that only contains the subset of adjacent nodes. Consensus is used to average the PDF of states in each node and the PDF of states received from neighbors. Given PDFs $p(\cdot)$ and $q(\cdot)$ and weight π , define the following information fusion and weighting operations:

$$p(x) \oplus q(x) = \frac{p(x)q(x)}{\int p(x)q(x)dx} \tag{36}$$

$$\pi \odot p(x) = \frac{p(x)^\pi}{\int p(x)^\pi dx} \tag{37}$$

The Kullback–Leibler average (KLA), which depends on relative entropy, is an average information definition of a PDF. The weighted KLA in PDF $\{p^i(\cdot)\}$ is defined as

$$\bar{p} = \arg \inf_{p(\cdot)} \sum_{i \in \mathcal{N}} \pi^i KL(p || p^i) \tag{38}$$

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

$$\lim_{l \rightarrow +\infty} p_l^i = \bar{p}(x), \forall i \in \mathcal{N} \tag{39}$$

where asymptotic PDF $\bar{p}(\cdot)$ represents the KLA with the same weight. The solution to this problem is the collective KLA: the weighted KLA is given by the normalized weighted geometric mean of PDF

$$\bar{p}(x) = \frac{\prod_{i \in \mathcal{N}} p^i(x)^{\pi^i}}{\int \prod_{i \in \mathcal{N}} p^i(x)^{\pi^i} dx} = \bigoplus_{i \in \mathcal{N}} (\pi^i \odot p^i(x)) \tag{40}$$

It can be calculated by updating local data in a distributed manner using a convex combination with data from neighbors

$$p_\ell^i(x) = \bigoplus_{j \in \mathcal{N}^i} (\pi^{i,j} \odot p_{\ell-1}^j(x)), \forall i \in \mathcal{N} \tag{41}$$

where ℓ is the number of consensus iteration steps, $\pi^{i,j}$ is the consensus weight satisfying $\pi^{i,j} \geq 0$, and $\sum_{j \in \mathcal{N}^i} \pi^{i,j} = 1$, and it is also the (i, j) component of the consensus matrix Π (if $j \notin \mathcal{N}^i$, $\pi^{i,j} = 0$). In addition, the initial value of iteration is $p_0^i(\cdot) = p^i(\cdot)$. Let $\pi_\ell^{i,j}$ be the (i, j) component of matrix Π^ℓ . Then,

$$p_\ell^i(x) = \bigoplus_{j \in \mathcal{N}} (\pi_\ell^{i,j} \odot p^j(x)), \forall i \in \mathcal{N}. \tag{42}$$

When $\pi^{i,j}$ is chosen so that matrix Π is primitive and doubly stochastic,

$$\lim_{\ell \rightarrow +\infty} \pi_\ell^{i,j} = \frac{1}{|\mathcal{N}|}, \forall i \in \mathcal{N} \tag{43}$$

Therefore, as the number of consensus steps increases, each local PDF tends to focus on the unweighted KLA.

For PDF $p(x^i) = N(x; m^i, P^i)$ with the Gaussian distribution, it can be proved that the probability density consensus algorithm can be simplified to an algebraic expression involving only its information vector $q^i = (P^i)^{-1}m^i$ and information matrix $\Omega^i = (P^i)^{-1}$:

$$\Omega_\ell^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} \Omega_{\ell-1}^j, i \in \mathcal{N}, \ell = 0, 1, \dots \tag{44}$$

$$q_\ell^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} q_{\ell-1}^j, i \in \mathcal{N}, \ell = 0, 1, \dots \tag{45}$$

This is the so-called CI consensus method.

In the above section, the multiple-distribution filter for a single sensor is given. To extend it to the distributed case, two important problems need to be solved: (1) for the mixed posterior distribution, which information should be transferred by the adjacent nodes and what strategy should be adopted to realize the collective KLA; and (2) since the mixed distribution is non-Gaussian, whether the CI consensus strategy for Gaussian distribution can be directly applied to the mixed distribution of Gaussian and student-t distributions.

For the first problem, the following consensus strategy is given: the distribution probability μ_k^r can be the first to obtain the consensus distribution probability, and then the

fused PDF $p^i(x) = \sum_{r=0}^1 \mu_k^r p^{i,r}(x)$ is given as the initial value, and then the consensus is run on the fused PDF. This strategy only needs to transfer the distribution probability and the fused PDF.

It should be noted that the above consensus method is for a continuous PDF. When the distribution probability is discrete, its distribution is a probability mass function (PMF). Given a PMF distribution $\mu^i (i \in \mathcal{N})$, the weighted KLA is defined as

$$\bar{\mu} = \arg \inf_{\mu^{(\cdot)}} \sum_{i \in \mathcal{N}} \pi^i KL(\mu || \mu^i) \tag{46}$$

For PMF μ, ν and weight $\pi > 0$, the following information fusion and weighting operations are defined:

$$\mu \oplus \nu = \eta = col(\eta^j)_{j \in M} \quad \text{with} \quad \eta^j = \frac{\mu^j \nu^j}{\sum_{h \in M} \mu^h \nu^h} \triangleq \mu^j \oplus \nu^j, \quad \forall j \in M, \tag{47}$$

$$\pi \odot \mu = \eta = col(\eta^j)_{j \in M} \quad \text{with} \quad \eta^j = \frac{(\mu^j)^\pi}{\sum_{h \in M} (\mu^h)^\pi} \triangleq \pi \odot \mu^j, \quad \forall j \in M \tag{48}$$

Then, the KLA probability mass function can be expressed as

$$\bar{\mu} = \bigoplus_{i \in \mathcal{N}} (\pi^i \odot \mu_i) \tag{49}$$

The collective fusion of the PMF can be obtained in a distributed manner:

$$\mu_\ell^i = \bigoplus_{j \in \mathcal{N}^i} (\pi^{i,j} \odot \mu_{\ell-1}^j) \tag{50}$$

where $\pi^{i,j} > 0$ is the consensus weight, and $\sum_{j \in \mathcal{N}^i} \pi^{i,j} = 1$.

Before answering the second question, let us first look at how to use a Gaussian distribution to approximate a student-t distribution. That is, how to find a scalar c to minimize the difference between $q(x) = N(x; \hat{x}, cP)$ and $p(x) = St(x; \hat{x}, P, \eta)$ under certain criteria. We know that when the dof of the student-t distribution tends to infinity, its distribution tends to the Gaussian distribution. Thus, we can obtain $q(x) = N(x; \hat{x}, cP) \approx St(x; \hat{x}, cP, \tilde{\eta} \rightarrow \infty)$. In this way, the problem becomes an approximation between two student-t distributions, so we can use the moment matching method to obtain the value of c , that is

$$c = \lim_{\tilde{\eta} \rightarrow \infty} \frac{\eta(\tilde{\eta} - 2)}{(\eta - 2)\tilde{\eta}} = \frac{\eta}{\eta - 2} \tag{51}$$

Therefore, $q(x) = N(x; \hat{x}, \frac{\eta}{\eta-2}P)$ can be used to approximate $p(x) = St(x; \hat{x}, P, \eta)$. We can see that the fusion process can be seen as the fusion of two Gaussian distributions. Thus, we can directly use the CI consensus method based on the Gaussian distribution.

3.3. The Distributed Multi-Distribution Filter

We have previously obtained the algorithm for a single sensor and the consensus strategy for multi-sensor mixed density. Now, we need to extend the results to the distributed case.

For local node i , the initial values $\hat{x}_0^{i,r}, P_0^{i,r}$, and μ_0^r are given for $r \in \mathcal{H}$. For the student-t filter, the common dof η is also given. When $k > 0$, start the following recursive process.

Step 1 Parallel filtering

For the Gaussian filter, we use Equations (9)–(15). For the student-t distribution filter, we make an approximation so that

$$\tilde{\eta}_{k-1}^i = \eta, c_k^i = \frac{(\eta + n_z)(\eta - 2)}{(\eta + n_z - 2)\eta} \tag{52}$$

$$\tilde{P}_{k-1}^1 = c_k^i P_{k-1}^1, \tilde{Q}_{k-1}^1 = c_k^i Q_{k-1}^1, \tilde{R}_{k-1}^1 = c_k^i R_{k-1}^1 \tag{53}$$

Then, Equations (17), (19) and (24) become

$$P_{k|k-1}^{i,1} = F_k \tilde{P}_{k-1}^{i,1} F_k^T + \tilde{Q}_{k-1}^{i,1} \tag{54}$$

$$S_k^{i,1} = H_k P_{k|k-1}^{i,1} (H_k)^T + \tilde{R}_k^{i,1} \tag{55}$$

$$\eta_k = \tilde{\eta}_{k-1} + n_z \tag{56}$$

Therefore, we can use Equations (16), (54), (18), (55), (20)–(23) and (56) for the student-t distribution filter. After that, we can obtain $\hat{x}_k^{i,r}, P_k^{i,r}, \hat{z}_k^{i,r}$, and $S_k^{i,r}$ for $r \in \mathcal{H}$ and η_k^i .

Step 2 Calculate distribution probability

Update probability

$$\mu_{k,0}^{i,r} = \frac{\Lambda_k^{i,r} \mu_{k-1}^{i,r}}{\sum_{j \in \mathcal{H}} \Lambda_k^{i,j} \mu_{k-1}^{i,j}} \tag{57}$$

where the likelihoods of the Gaussian and student-t filter $\Lambda_k^{i,0}$ and $\Lambda_k^{i,1}$ can be obtained by Equations (27) and (28).

Step 3 Consensus on distribution probability

For L-step consensus,

$$\mu_{k,\ell}^{i,r} = \bigoplus_{j \in \mathcal{N}^i} (\pi^{i,j} \odot \mu_{k,\ell-1}^{i,r}), \quad \ell = 1, \dots, L, r \in \mathcal{H} \tag{58}$$

$$\mu_k^{i,r} = \mu_{k,L}^{i,r} \tag{59}$$

where $\ell = 1, 2, \dots, L$ is the number of consensus steps, $\pi^{i,j}$ is the consensus weight, $\pi^{i,j} \geq 0$, and $\sum_{j \in \mathcal{N}^i} \pi^{i,j} = 1$.

Step 4 Fuse the mixed PDF

$$\hat{x}_{k,0}^i = \sum_{r \in \mathcal{H}} \mu_k^{i,r} \hat{x}_k^{i,r} \tag{60}$$

$$\tilde{P}_k^{i,1} = \frac{\eta_k}{\eta_k - 2} P_k^{i,1} \tag{61}$$

$$P_{k,0}^i = \mu_k^{i,0} P_k^{i,0} + \mu_k^{i,1} \tilde{P}_k^{i,1} + \sum_{r \in \mathcal{H}} \mu_k^{i,r} (\hat{x}_k^{i,r} - \hat{x}_{k,0}^i)(\hat{x}_k^{i,r} - \hat{x}_{k,0}^i)^T \tag{62}$$

Step 5 Consensus on fused PDF

$\Omega_{k,0}^i = (P_{k,0}^i)^{-1}, q_{k,0}^i = \Omega_{k,0}^i x_{k,0}^i$
 For L-step consensus,

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

$$q_{k,\ell}^i = \sum_{j \in \mathcal{N}^i} \pi^{i,j} q_{k,\ell-1}^j, \quad \ell = 1, \dots, L \tag{64}$$

Step 6 Reinitialization

$$\hat{x}_k^{i,0} = (\Omega_{k,L}^i)^{-1} q_{k,L}^i, P_k^{i,0} = (\Omega_{k,L}^i)^{-1} \tag{65}$$

$$\hat{x}_k^{i,1} = (\Omega_{k,L}^i)^{-1} q_{k,L}^i P_k^{i,1} = \frac{\eta_k - 2}{\eta_k} (\Omega_{k,L}^i)^{-1} \tag{66}$$

The workflow of the proposed algorithm is shown in Figure 4, and the pseudocode is summarized in Algorithm 1.

Algorithm 1: Distributed consensus multi-distribution filter (DCMDF)

Given the initial values $\hat{x}_0^{i,r}$, $P_0^{i,r}$ and μ_0^r for $r \in \mathcal{H}$ and the common dof η , for each time step k at every node i , start the following recursive process:

Step1 Parallel filtering:

Gaussian filtering: Calculate $\hat{x}_k^{0,i}$, $P_k^{0,i}$, $S_k^{0,i}$ by Equations (9)–(15)

Student-t filtering: Calculate $\hat{x}_k^{1,i}$, $P_k^{1,i}$, $S_k^{1,i}$ by Equations (16)–(22) and (54), (55)

$$\eta_k = \tilde{\eta}_{k-1} + n_z$$

Step2 Calculate distribution probability:

Calculate $\mu_{k,0}^{i,r}$ according to (57)

Step3 Consensus on distribution probability:

for $\ell = 1$ to L

 Calculate $\mu_{k,\ell}^{i,r}$ according to (58)

end for

$$\mu_k^{i,r} = \mu_{k,L}^{i,r}$$

Step4 Fuse the mixed PDF:

Calculate $\hat{x}_{k,0}^i$ and $P_{k,0}^i$ according to (60)–(62)

Step5 Consensus on fused PDF:

$$\Omega_{k,0}^i = (P_{k,0}^i)^{-1}, q_{k,0}^i = \Omega_{k,0}^i \hat{x}_{k,0}^i$$

for $\ell = 1$ to L

 Calculate $\Omega_{k,\ell}^i$ and $q_{k,\ell}^i$ according to (63) and (64)

end for

Step6 Reinitialization:

$$\hat{x}_k^{i,0} = (\Omega_{k,L}^i)^{-1} q_{k,L}^i P_k^{i,0} = (\Omega_{k,L}^i)^{-1}$$

$$\hat{x}_k^{i,1} = (\Omega_{k,L}^i)^{-1} q_{k,L}^i P_k^{i,1} = \frac{\eta_k - 2}{\eta_k} (\Omega_{k,L}^i)^{-1}$$

Return: $\hat{x}_k^{i,0}$, $P_k^{i,0}$, $\hat{x}_k^{i,1}$ and $P_k^{i,1}$

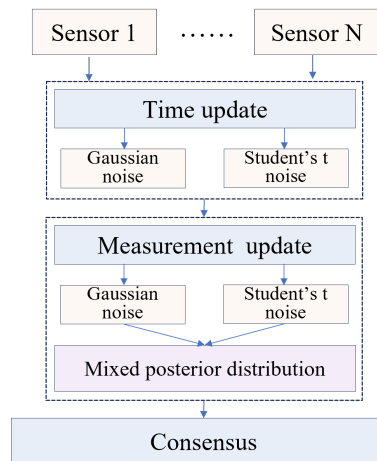


Figure 4. The workflow of the proposed algorithm.

4. Numerical Simulation

Consider a linear dynamic system in which the state contains $x = [p_x, \dot{p}_x, p_y, \dot{p}_y]^T$:

$$F_k = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}, Q_k = G_k \Delta G_k^T \tag{67}$$

$$G_k = \begin{bmatrix} T^2/2 & T & 0 & 0 \\ 0 & 0 & T^2/2 & T \end{bmatrix}^T \tag{68}$$

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

There are 20 sensor nodes in the sensor network, and their graphical topology is shown in Figure 5. The measurement model is

$$H_k^i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{69}$$

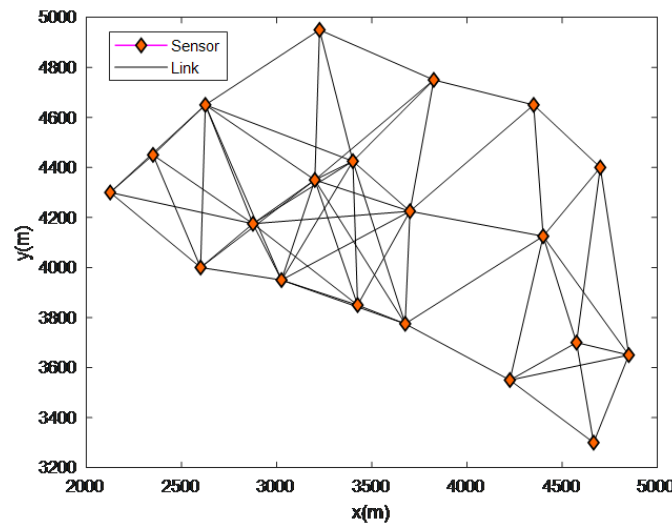


Figure 5. Topology of sensor network.

The standard measurement noise variance is $R = \text{diag}([(15\text{ m})^2, (15\text{ m})^2])$.

Assume that the initial true state is

$$x_0 = [2600\text{ m}, 20\text{ m/s}, 3800\text{ m}, 10\text{ m/s}]^T \tag{70}$$

In the simulation, the initial state of the filter is randomly selected from $N(x_0, P_0)$, where

$$P_0 = \text{diag}([50^2\text{ m}^2, 5^2\text{ m}^2/\text{s}^2, 50^2\text{ m}^2, 5^2\text{ m}^2/\text{s}^2]) \tag{71}$$

Process noise and measurement noise with outliers are generated by the following model [20,23]:

$$w_k \sim \begin{cases} N(0, Q) & \text{with probability } 1 - p_o \\ N(0, 100Q) & \text{with probability } p_o \end{cases} \tag{72}$$

$$v_k^i \sim \begin{cases} N(0, R) & \text{with probability } 1 - p_o \\ N(0, 100R) & \text{with probability } p_o \end{cases} \tag{73}$$

where p_o is the probability of measuring outliers. It should be noted that this model for heavy-tailed noise is widely used.

In this section, the following three methods are compared: (1) distributed consensus Kalman Filter [14], referred to as DCKF; (2) distributed consensus student-t filter, proposed

in [23], referred to as DCSTF; (3) the proposed algorithm, referred to as DCMDF; and (4) the multiple-model method with two Gaussian distributions, referred to as DCKFIMM.

The number of consensus steps was set to $L = 3$, and the consensus weight of the sensor node was 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 dof of the student-t distribution in DCSTF and DCMDF was $\eta = 10$. One model of DCKFIMM is the same as the DCKF, and another model has the following parameters: the measurement noise variance is $100R$, and the process noise variance is $100Q$. The simulation results were obtained through 100 Monte Carlo simulations and evaluated using the root mean square errors (RMSEs) of the position and velocity.

The occurrence probability of outliers was set to $p_o = 0$, which means that the noises are Gaussian. The RMSEs of the position and velocity of the four algorithms are shown in Figure 6 and 7. It can be found that under the condition of a Gaussian distribution, the performance of the DCMDF algorithm is almost the same as that of DCKF, while the error of DCSTF is relatively large. The performance of the DCKFIMM algorithm is better than the DCSTF and little worse than the DCMDF. This result shows that Gaussian models are better than the single student-t model in dealing with estimation problems in normal noise environments. At the same time, it is worth noting that DCMDF adopts a mixed model of Gaussian distribution and student-t distribution; therefore, it can obtain a higher estimation accuracy through model adaptation.

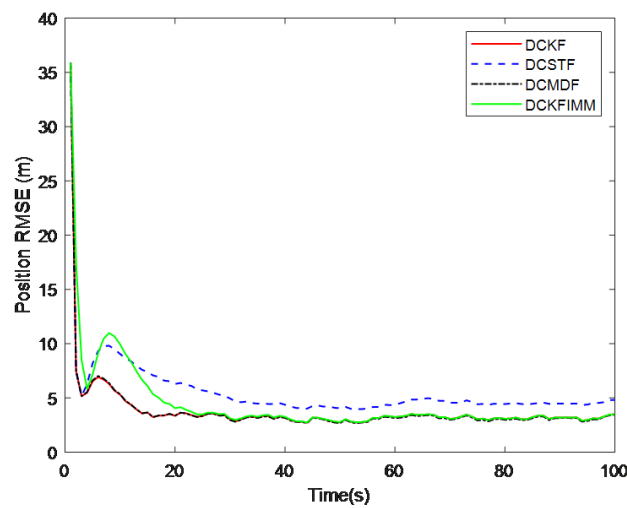


Figure 6. Position root mean square errors of different algorithms.

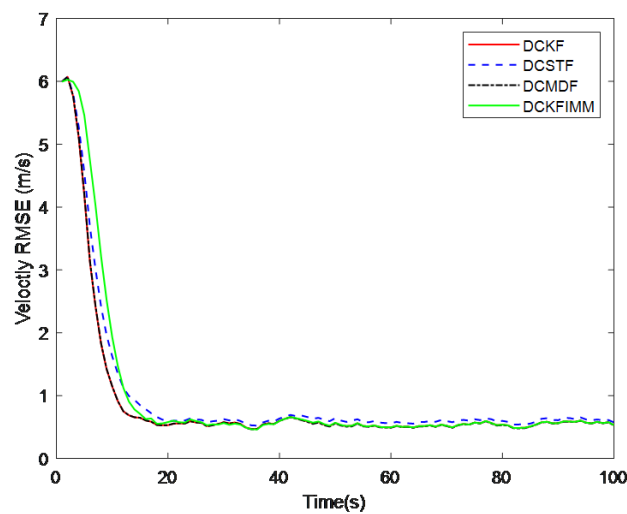


Figure 7. Velocity root mean square errors of different algorithms.

The occurrence probability of outliers was set to $p_o = 0.3$. Figure 8 presents the true and estimated trajectories for one simulation. Figures 9 and 10 show the RMSEs of the position and velocity of the three algorithms. It can be found that DCKF has the worst performance, which is caused by the natural defects of the Gaussian distribution. Although it can achieve good performance in a normal noise environment, it struggles to dealing with the estimation in the presence of outliers. The DCKFIMM has a better performance than DCKF; however, the performance of the DCKDIMM is worse than those of DCSTF and DCMDF because DCSTF and DCMDF adopt the student-t distribution, which is more in line with the distribution of noise while outliers occur. It is worth noting that DCMDF has the best performance because the algorithm uses a hybrid distribution model, which can be adaptive for the model in the estimation process, making the algorithm more robust and accurate.

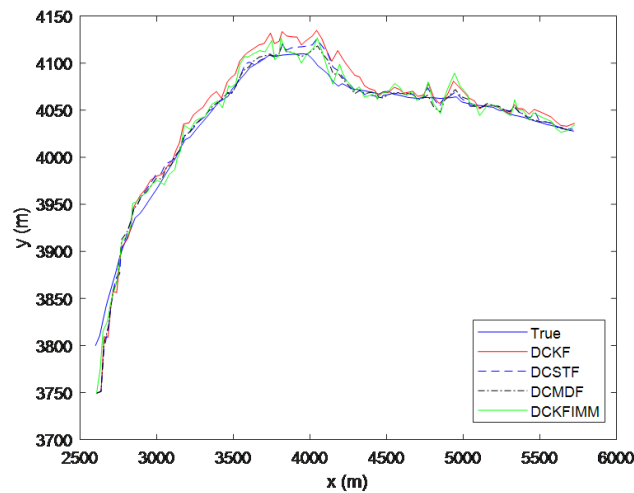


Figure 8. The true and estimated trajectories for one simulation.

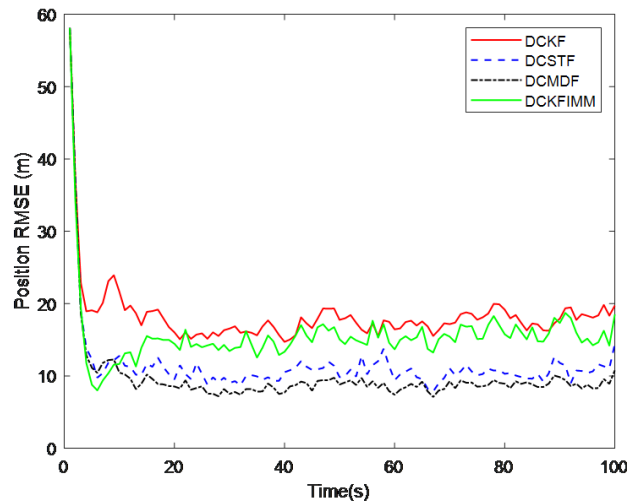


Figure 9. Position root mean square errors of different algorithms.

Tables 3 and 4 show the simulation results of the three algorithms under different outlier probabilities. It can be found that the RMSEs of the four algorithms will increase with the increase in outlier probabilities. However, the estimation performance of DCMDF always shows a relatively better estimation performance, and the estimation accuracy of the proposed algorithm improved by at least 20% compared to the other algorithms when $p_o \neq 0$, which further proves that DCMDF has a good performance in robustness and estimation accuracy.

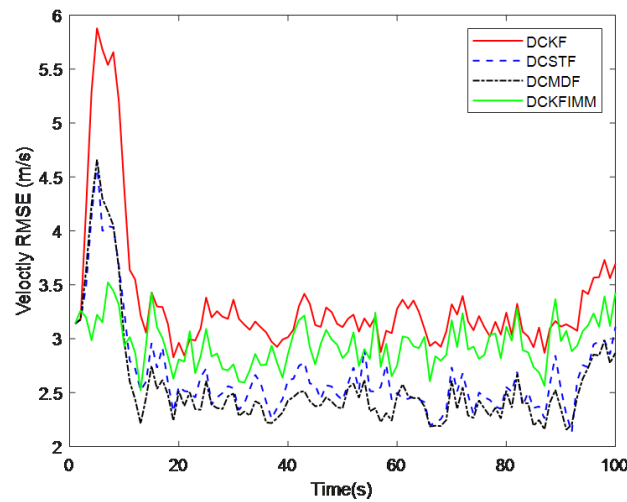


Figure 10. Velocity root mean square errors of different algorithms.

Table 3. Root mean square errors of position under different outlier probabilities.

p_0	DCKF	DCSTF	DC MDF	DCKFIMM
0	3.8339	5.5194	3.7899	4.4447
0.1	11.4991	8.9549	7.0405	9.0710
0.2	15.3950	9.4373	8.1685	11.5212
0.3	18.2466	11.4850	9.6818	14.0321
0.4	20.9295	14.8262	12.2020	18.3049

Table 4. Root mean square errors of velocity under different outlier probabilities.

p_0	DCKF	DCSTF	DC MDF	DCKFIMM
0	0.9405	1.0602	0.9376	1.0908
0.1	2.4378	2.2842	2.1637	2.2934
0.2	3.0931	2.6195	2.5565	2.3951
0.3	3.5050	2.9143	2.7751	3.0377
0.4	3.9130	3.3194	3.03984	3.4926

To further test the performance of the algorithm, we reduced the number of nodes and density of network. The topology of a sensor network with 15 nodes is shown in Figure 11. Since the nodes and density of network were reduced, we increased the number of consensus steps to 6, the RMSEs of the position and velocity are shown in Tables 5 and 6. It can be seen that the proposed algorithm still has better performances in estimation accuracy.

In order to handle the estimation problem with heavy-tailed noise, in this study, we made two hypotheses for the model noise, which include the Gaussian distribution and student-t distribution. The reason why the student-t distribution was chosen to represent the heavy-tailed noise distribution is discussed in detail in [23], which is briefly reviewed here. By adjusting the dof, the student-t distribution can tend toward a Gaussian distribution and Cauchy distribution, which have good flexibility. Similar to the traditional multiple-model algorithm, we calculate the state estimation results based on two different hypothetical noises. Then, the two results are fused based on different weights. Therefore, the proposed algorithm can balance the robustness and estimation accuracy.

The computation and communication burden are important issues in distributed-state estimation. Compared with the traditional CI algorithm, the computation of the proposed algorithm is mainly in the step of local estimation. The amount of computation has doubled, but the strategy of parallel computing can be adopted. The other step is to calculate the

distribution probability of consensus, which has a small increase in calculation. The increase in communication also comes from this step. It should be noted that the quantities that we are concerned with are the state and its error matrix. This step can also be omitted in the application. Then, the increases in computation and communication from the distribution probability calculation and consensus will be eliminated, which can further improve the computational efficiency and reduce the communication burden.

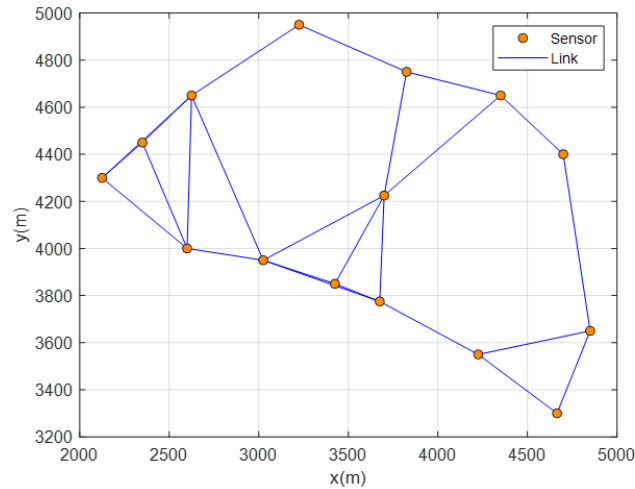


Figure 11. Topology of sensor network with 15 nodes.

Table 5. Root mean square errors of position under different outlier probabilities.

p_0	DCKF	DCSTF	DCMDF	DCKFIMM
0	3.7362	4.7315	3.7021	4.1563
0.1	11.5420	8.4957	6.9921	9.8612
0.2	15.2641	9.5924	8.0813	11.6715
0.3	18.8536	11.4042	9.7624	15.6121
0.4	19.7912	13.8126	11.4620	17.9736

Table 6. Root mean square errors of velocity under different outlier probabilities.

p_0	DCKF	DCSTF	DCMDF	DCKFIMM
0	0.7623	0.8325	0.7546	0.8216
0.1	1.9820	1.7741	2.0831	1.8742
0.2	3.0362	2.6135	2.5272	2.6718
0.3	3.4931	2.8635	2.7452	3.1803
0.4	3.8842	3.3146	3.0653	3.6943

5. Conclusions and Outlook

For the problem of distributed state estimation in which both process noise and measurement noise are heavy-tailed, a distributed consistent multi-distribution state estimation algorithm based on the Gaussian distribution and student-t distribution parallel filter is proposed in this paper, with the following steps:

- (1) The system model is established as two models based on the Gaussian distribution and student-t distribution, and each model is assigned a filter based on the distribution. An algorithm based on single-sensor observation is derived, and a combined posterior distribution based on a mixed posterior distribution is presented. To solve the problem of the filter converging to the Gaussian distribution quickly, the moment matching method is used to approximate the filter to keep the heavy-tailed characteristics.

- (3) Aiming at the consensus problem of mixed probability density, a mixed strategy based on CI consensus is proposed: firstly, the consensus on the discrete distributed probability PMF is carried out to obtain the KLA PMF; then, the posterior of the multi-distributions is combined using the probability of the distribution to obtain the posterior of the combination; finally, the CI method is used to obtain the consensus of the posterior of the combination. Based on the above hybrid probability consensus strategy and single-sensor algorithm, the recursive steps of distributed consensus multi-distribution filtering are presented.
- (3) The simulation results show that the proposed algorithm can achieve good results in both Gaussian noise and heavy-tailed noise scenarios.

Although the method proposed this paper deals with state estimation with heavy-tailed noise, it is still a type of rough processing method. The algorithm can ensure a low-bound estimation accuracy, but there is still potential for improvement.

For example, in target tracking, the process transfer model is time-variant. The two types of noise hypotheses in this paper can model the uncertainty of a process transfer, but this modeling method is still not precise enough. In a broad sense, the algorithm proposed in this paper can be regarded as a special case of the multi-model method. Therefore, the concepts of probability transfer and input interaction of the multi-model method can also be applied to the algorithm proposed in this paper. How to integrate these technologies is one of the problems to be solved in the future.

Author Contributions: Conceptualization, G.-N.C. and W.-X.F.; methodology, G.-N.C. and P.D.; software, Guan-nan Chan, T.C. and L.-Y.S.; validation, G.-N.C., W.-X.F. and T.C.; formal analysis, G.-N.C. and W.-X.F.; investigation, W.-X.F.; resources, P.D.; data curation, W.-X.F.; writing—original draft preparation, G.-N.C. and L.-Y.S.; writing—review and editing, G.-N.C. and T.C.; visualization, G.-N.C. and L.-Y.S.; supervision, P.D.; project administration, W.-X.F.; funding acquisition, P.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was jointly supported by the National Natural Science Foundation of China (No. 61803260).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

Acronyms and Nomenclatures

Notations	Definitions	Notations	Definitions
CE	consensus on estimation	CM	consensus on measurement
CI	consensus on information	HCMCI	hybrid CM-CI
dof	degrees of freedom	PDF	probability density function
KLA	Kullback–Leibler average	KLD	Kullback–Leibler divergence
DCKF	distributed consensus Kalman Filter	DCSTF	distributed consensus student-t filter
DCMDF	distributed consensus multiple-distribution filter	DCKFIMM	distributed multiple-model method with two Gaussian distributions
RMSE	root mean square error	PMF	probability mass function

References

- Ding, D.; Han, Q.L.; Wang, Z.; Ge, X. A survey on model-based distributed control and filtering for industrial cyber-physical systems. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2483–2499. [[CrossRef](#)]
- He, S.; Shin, H.; Xu, S.; Tsourdos, A. Distributed estimation over a low-cost sensor network: A review of state-of-the-art. *Inform. Fusion* **2020**, *54*, 21–43. [[CrossRef](#)]
- Huang, S.; Li, Y.; Wu, J. Distributed state estimation for linear time-invariant dynamical systems: A review of theories and algorithms. *Chin. J. Aeronaut.* **2022**, *35*, 1–17. [[CrossRef](#)]
- Sirocchi, C.; Bogliolo, A. Community-Based Gossip Algorithm for Distributed Averaging. In Proceedings of the IFIP International Conference on Distributed Applications and Interoperable Systems, Lisbon, Portugal, 19–23 June 2023.

5. Loizou, N.; Richtarik, P. Revisiting randomized gossip algorithms: General framework, convergence rates and novel block and accelerated protocols. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2483–2499. [[CrossRef](#)]
6. Cai, P.; Wang, S.; Chen, Y. Diffusion Mixture Minimum Total Error Entropy Adaptive Filtering Algorithm and Its Performance Analysis. *IEEE Trans. Signal Inf. Process. Over Netw.* **2023**, *9*, 397–411. [[CrossRef](#)]
7. Liang, Y.; Li, Y.; Chen, Y.; Sheng, A. Event-triggered diffusion nonlinear estimation for sensor networks with unknown cross-correlations. *Syst. Control Lett.* **2023**, *175*, 105506. [[CrossRef](#)]
8. Olfati-Saber, R. Distributed Kalman filter with embedded consensus filters. In Proceedings of the 44th IEEE Conference on Decision and Control, Seville, Spain, 15 December 2005.
9. Olfati-Saber, R.; Murray, R.M. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control* **2004**, *49*, 1520–1533. [[CrossRef](#)]
10. Battistelli, G.; Chisci, L.; Fantacci, C. Parallel consensus on likelihoods and priors for networked nonlinear filtering. *IEEE Signal Process. Lett.* **2014**, *21*, 787–791. [[CrossRef](#)]
11. Battistelli, G.; Chisci, L.; Mugnai, G.; Farina, A.; Graziano, A. Consensus-based linear and nonlinear filtering. *IEEE Trans. Autom. Control* **2015**, *60*, 1410–1415. [[CrossRef](#)]
12. Olfati-Saber, R. Distributed Kalman filtering for sensor networks. In Proceedings of the 2007 46th IEEE Conference on Decision and Control, New Orleans, LA, USA, 12–14 December 2007.
13. Battistelli, G.; Chisci, L.; Morrocchi, S.; Papi, F. An information-theoretic approach to distributed state estimation. *IFAC Proc. Vol.* **2011**, *44*, 12477–12482. [[CrossRef](#)]
14. Battistelli, G.; Chisci, L. Kullback–Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. *Automatica* **2014**, *50*, 707–718. [[CrossRef](#)]
15. He, X.; Xue, W.; Fang, H. Consistent distributed state estimation with global observability over sensor network. *Automatica* **2018**, *92*, 162–172. [[CrossRef](#)]
16. Qu, H.; Yang, F.; Han, Q.; Zhang, Y. Distributed H_∞ consensus filtering for attitude tracking using ground-based radars. *IEEE Trans. Cybern.* **2021**, *51*, 3767–3778. [[CrossRef](#)] [[PubMed](#)]
17. Rego, F.C.; Pu, Y.; Alessandretti, A.; Aguiar, A.P.; Pascoal, A.M.; Jones, C.N. A distributed Luenberger observer for linear state feedback systems with quantized and rate-limited communications. *IEEE Trans. Autom. Control* **2021**, *66*, 3922–3937. [[CrossRef](#)]
18. Ting, J.A.; Theodorou, E.; Schaal, S. A Kalman filter for robust outlier detection. In Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 29 October–2 November 2007.
19. Fu, H.; Cheng, Y. A Novel Robust Kalman Filter Based on Switching Gaussian-Heavy-Tailed Distribution. *IEEE Trans. Circuits Syst. II Express Briefs* **2022**, *69*, 3012–3016. [[CrossRef](#)]
20. 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.* **2019**, *55*, 1545–1554.
21. Huang, Y.; Zhang, Y.; Xu, B.; Wu, Z.; Chambers, J. A New Outlier-Robust Student’s t Based Gaussian Approximate Filter for Cooperative Localization. *IEEE-ASME Trans. Mechatron.* **2017**, *22*, 2380–2386. [[CrossRef](#)]
22. Dong, P.; Jing, Z.; Shen, K.; Li, M. A distributed consensus filter for sensor networks with heavy-tailed measurement noise. *Sci. China Inform. Sci.* **2018**, *61*, 119201. [[CrossRef](#)]
23. Wang, J.; Dong, P.; Shen, K.; Song, X.; Wang, X. Distributed consensus student-t filter for sensor networks with heavy-tailed process and measurement noises. *IEEE Access* **2020**, *8*, 167865–167874. [[CrossRef](#)]
24. Huang, Y.; Zhang, Y.; Zhao, Y.; Chambers, J.A. A novel robust Gaussian–student’s t mixture distribution based Kalman filter. *IEEE Trans. Signal Process.* **2019**, *67*, 3606–3620. [[CrossRef](#)]
25. Roth, M.; Ozkan, 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, Canada, 26–31 May 2013.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.