Signal Processing 200 (2022) 108678

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Tightly coupled distributed Kalman filter under non-Gaussian noises

Yangqing Fu[a], Ming Sun[b], Yue Gao[c,*]

[a] Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China
[b] Department of Automation, Shanghai Jiao Tong University, Shanghai, China
[c] MoE Key Lab of Artificial Intelligence, Department of Automation, Shanghai Jiao Tong University, Shanghai, China

ARTICLE INFO

Article history:
Received 6 May 2022
Revised 26 June 2022
Accepted 4 July 2022
Available online 5 July 2022

Keywords:
Distributed Kalman filter
State estimation
Covariance intersection
Gaussian mixture model

ABSTRACT

Distributed Kalman filter (DKF) is an effective method to solve the problem of state estimation in multi-sensor systems. However, the accuracy of conventional DKF may be influenced by the non-Gaussian noises and the accumulated errors in local Kalman filters (LKFs). To encounter the above challenges, a tightly coupled DKF with covariance intersection is proposed. In our method, using the computation outputs of LKFs, a Gaussian Mixture Model is formulated to address the influence of non-Gaussian noises. In order to reduce the cumulative errors from LKFs, covariance intersection fusion is utilized. Furthermore, an index-Huber function is designed to reduce the impact of large covariance generated by the LKFs. Several simulation and real-world experiments are conducted to show the effectiveness of our methods. The method we proposed outperforms three other DKF algorithms in the metrics of RMSE and cumulative error. In addition, a real-world multi-sensor state estimation experiment is conducted on a hexapod robot.

1. Introduction

Multi-sensor fusion has received much attention in recent years due to its wide applications in complex scenarios, such as GPS position, mobile robot navigation, and target tracking [1–3]. Centralized state estimation method is limited by the massive resources consumption in computation and communication in the multi-sensor systems [4]. To solve this challenge, distributed Kalman filter (DKF) is proposed for complex systems with multiple sensor nodes and multiple agents [5]. For a conventional DKF, there is an individual Kalman filter on each sensor node, and the final state fusion is completed by centralized fusion [6,7]. Distributed structure can provide local information of sensor nodes and prevent the influence of outliers on global estimation [8,9]. The approach to fuse the estimation results of multiple Kalman filters is an important issue in DKF. The traditional method is based on the covariance matrix to calculate the fusion results [10], while some recent methods utilize a centralized Kalman filter for fusion based on the embedded average [11]. For variants of DKF, such as distributed unscented Kalman filter, a weighted average consensus-based algorithm is introduced in [12]. For a realistic system which is generally a nonlinear stochastic system, the noises of the sensor measurements may not follow Gaussian distributions, which violates the assumption of traditional Kalman filter and causes modeling inconsistency with the actual system [13,14].

Recently, some researchers have proposed a series of methods to handle non-Gaussian noises for DKF. Chen et al. [15] proposed a weighted optimization-based distributed Kalman filter algorithm which creates a new cost function to estimate the model noise covariance, and the experiment results demonstrate its effectiveness to improves the estimation accuracy in non-linear and non-Gaussian tracking problems. In addition, a distributed Gaussian sum filter was proposed in [16] to track multiple moving targets in a non-Gaussian sensor network. In [2], the Gaussian mixture model (GMM) is combined with distributed unscented Kalman filter, which can overcome the limitation of Kalman filter on nonlinear and non-Gaussian systems. Compared with a single Gaussian distribution, GMM can more accurately reflect the data distribution in the real world [17,18]. When the data completely follow a Gaussian distribution, GMM will degenerate into a single Gaussian distribution. In the presence of multi-sensor error registration, the Gaussian mixture implementation can give a closed-form solution under the linear Gaussian assumptions on the bias dynamics [19]. Since the solution of GMM is generally solved by Expectation-Maximization (EM) algorithm, the fusion results from GMM are not guaranteed to be optimal. The accuracy of centralized fusion is directly influenced by the weights and the covariances of the distributions, but current works have not taken this problem into consideration. In this paper, we analyze the relationship between the

* Corresponding author.
E-mail addresses: frank79110@sjtu.edu.cn (Y. Fu), mingsun@sjtu.edu.cn (M. Sun), yuegao@sjtu.edu.cn (Y. Gao).

https://doi.org/10.1016/j.sigpro.2022.108678
0165-1684/

weight and the covariance of the LKF, and propose a novel index-Huber function to reduce the weight of distribution with large covariance.

Since the cumulative error can reduce the fusion accuracy and lead to the divergence of the filter, DKF algorithms should consider the influence of the cumulative errors from local filters on the fusion result [20]. Considering that the local and central state estimates for all nodes are the same, Olfati et al. proposed consensus-based DKF, which uses the consensus result to update local filters [21]. In addition, covariance intersection is a commonly used fusion method for state correction of the local filter. Distributed covariance intersection is discussed in [22], and experiments show that sequential covariance intersection fusion approach improves the accuracy of distributed filter. And Based on previous research, our proposed method utilizes covariance intersection to construct a tightly coupled DKF structure.

In this paper, a novel tightly coupled DKF algorithm utilizing covariance intersection is proposed to correct local filters. Different from previous work, we focus more on the effect of GMM weights on fusion accuracy and the cumulative error of local filter. In our method, GMM is adopted to handle non-Gaussian priors. The local filters in our framework may be Kalman filter and its variants, including extended Kalman filter (EKF), unscented Kalman filter (UKF), etc. The proposed method is adapted to solve the problem of state estimation and information fusion for systems with multiple sensor nodes under non-Gaussian noise, such as power network, vehicle tracking, and robotics. The main contributions of this paper are as follows:

(i) GMM is utilized to fuse the results of local Kalman filters (LKFs), which are represented as individual Gaussian distributions. Furthermore, a closed-form solution for solving GMM is provided. And a tightly coupled DKF framework with covariance intersection is developed.
(ii) The relationship between the local covariance and the GMM weights is analyzed. And an index-Huber function is designed to dynamically reduce the effect of large covariance from LKFs on the centralized fusion.
(iii) A simulation experiment and a real-wold experiment on a hexapod robot with multiple sensors are conducted to evaluate the proposed method.

The remainder of this paper is organized as follows: Section 2 introduces the preliminaries, which describes DKF and GMM. In Section 3, the framework of our method and the specific methodology are presented. Then Section 4 provides experiment results of a Monte-Carlo simulation example as well as a real-world multi-sensor system on a hexapod robot. Finally the conclusions and discussions are conducted in Section 5.

2. Preliminary study

2.1. Distributed Kalman filter

The definitions of notations in this paper are listed in Table 1. Consider a system with K sensors. Assuming the measurement of the kth sensor at timestep t is y_t^k and the true state of the system is $x_t \in \mathbb{R}^N$. The distributed dynamic system can be described by

$$x_t = \phi_t x_{t-1} + c_t \tag{1}$$

$$y_t^k = H_t^k x_t + v_t^k \tag{2}$$

where the process noise $c_t \sim \mathcal{N}(0, Q)$. And ϕ_t is the state transition of the system at timestep t. Assume that there are K nodes acquiring measurements from different sensors, the local measurement updating is described by Eq. (2). And $v_t^k \sim \mathcal{N}\left(0, R_t^k\right)$ is the measurement noise of the kth sensor node.

The state estimation results at each node are given by LKFs. The update and prediction equations of kth LKF based on its measurement are expressed as

$$\hat{K}_t^k = \hat{P}_{t/t}^k H_t^k (H_t^k \hat{P}_{t/t}^k H_t^{k^T} + \hat{R}_t^k)^{-1} \tag{3}$$

$$\hat{x}_{t/t}^k = \hat{x}_{t/t-1}^k + \hat{K}_t^k \left(y_t^k - H_t^k \hat{x}_{t/t-1}^k\right) \tag{4}$$

$$\hat{P}_{t/t}^k = \hat{\Phi}_t^k \hat{P}_{t/t-1}^k \hat{\Phi}_t^{k^T} + \hat{Q}_t^k \tag{5}$$

where $\hat{K}_t^k$ is the Kalman gain of the kth sensor node and $\hat{\Phi}_t^k$ is the state transition. The output of the Kalman filter is actually the optimal estimation of the predicted distribution and the measurement distribution [23]. In addition, the estimation result belongs to Gaussian distribution $\mathcal{N}_t^k(\hat{x}_{t/t}^k, \hat{P}_{t/t}^k)$. The covariance matrix $\hat{P}_t^k$ of the kth-LKF can be calculated by Eq. (5). The centralized covariance fusion can be described as

$$\hat{P}_{t/t}^{-1} = \hat{P}_{t/t-1}^{-1} + \sum_{i=1}^{K} \left(\hat{P}_{t/t}^{i^{-1}} - \hat{P}_{t/t-1}^{i^{-1}}\right) \tag{6}$$

2.2. Gaussian mixture model

A GMM is a parametric probability model represented as a weighted sum of multiple single Gaussian models [24]. Multivariate GMM composed of K component Gaussian densities is given by the following equation:

$$p(\mathbf{x} \mid \theta) = \sum_{i=1}^{K} w_i g(\mathbf{x} \mid \theta_k) \tag{7}$$

where $x = (x_1, x_2, \ldots, x_n)^T$ is a n-dimensional vector, which refers to n-dimensional input data; $\theta = (\mu_m, \Sigma_m)$ represents the mean and covariance of GMM, and $\theta_k = (\mu_k, \Sigma_k)$ represents the mean and covariance of kth Gaussian distribution with mixture weight w_k. The probability density function of the kth component Gaussian distribution is

$$g(x_k; \mu_k, \Sigma_k) = \frac{\exp\left(-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)\right)}{(2\pi)^{n/2}|\Sigma_k|^{1/2}} \tag{8}$$

Compared with an individual Gaussian distribution, GMM represents data distribution using multiple stochastic processes, hence it is more capable of describing the complicated distributions in real-world scenarios.

ExpectationMaximization (EM) algorithm [25] is a classical and effective method to train a GMM, i.e., to determine the parameters $w_k, \theta_k (k = 1, \ldots, K)$. The main idea of EM algorithm is to maximize the likelihood of the data set and it needs multiple iterations. However, state estimation problems are usually sensitive to the computational complexity due to the real-time requirement. Therefore, in our method, LKF directly computes the mean $\hat{x}_{t/t}^k$ and the covariance $\hat{P}_t^k$, and the maximum likelihood estimation only computes the parameter w_t^k. In this way, GMM can be solved quickly without tedious iterations, which reduces the computational cost.

3. Tightly coupled DKF under non-Gaussian noises

3.1. Design of the framework

The framework of the proposed method are shown as Fig. 1. According to the prior knowledge, the multivariate state variable $\hat{x}_{t-1/t-1}^k$ follows the multivariate Gaussian distribution with covariance matrix $\hat{P}_{t-1/t-1}^k$. Based on local dynamics assumption

Table 1
Notation definitions.

Notation	Description
t	timestep
k	kth sensor node
$\hat{x}^k_{t/t-1}$	state of kth sensor node after prediction at timestep t
$\hat{x}^k_{t/t}$	state of kth sensor node after updating measurements at timestep t
$\hat{P}^k_{t/t-1}$	covariance matrix of kth sensor node after prediction at timestep t
$\hat{P}^k_{t/t}$	covariance matrix of the kth sensor after updating measurements at timestep t
$\hat{x}^m_t$	fused state based on GMM at timestep t
$\hat{P}^m_t$	fused covariance matrix based on GMM at timestep t
w^k_t	weight of kth Gaussian distribution at timestep t
λ^k_t	weight of kth LKF covariance matrix at timestep t

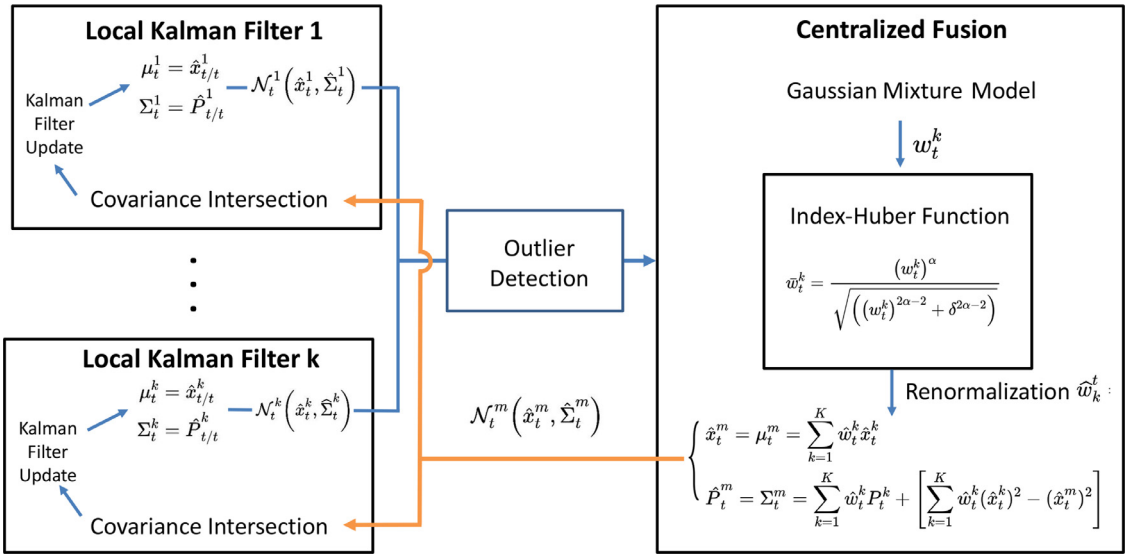


Fig. 1. The framework of the proposed tightly coupled DKF. The GMM is composed of the local Kalman filter results after outlier detection and the weights are computed through maximum likelihood estimation. The covariance intersection is utilized to minimize the trace of covariance matrix in the local Kalman filter.

$\hat{\Phi}^k_t$ with local process noise $c^k_{t-1} \sim \mathcal{N}(0, \hat{Q}^k_t)$, the prediction states $\hat{x}^k_{t/t-1}$ and corresponding covariance matrix $\hat{P}^k_{t/t-1}$ can be computed by Eqs. (4) and (5). The local information of the kth sensor node provides the posterior distribution y^k_t with measurements noise $v^k_t \sim \mathcal{N}(0, \hat{R}^k_t)$ for the kth LKF, and the optimal estimation result $\hat{x}^k_{t/t}$ with covariance matrix $\hat{P}^k_{t/t}$ can be computed. The estimation result follows the multivariate Gaussian distribution $\mathcal{N}^k_t(\hat{x}^k_{t/t}, \hat{P}^k_{t/t})$. Since outliers can cause a non-negligible impact on weight computation of GMM, the outlier detection algorithm is used to reject abnormal distribution. Then maximum likelihood estimation is used to compute the mixture weight w^k_t for each LKF. In order to reduce the influence of the distribution with large covariance, an Index-Huber function is designed to decrease weights of the distributions with large covariance and re-normalize weights $\hat{w}^k_t$, where $\sum_{k=1}^K \hat{w}^k_t$. At the end of this timestep, the mean $\hat{x}^m_t$ and covariance $\hat{P}^m_t$ of GMM is calculated and the LKF is updated with covariance intersection. The covariance intersection can further improve the estimation accuracy of the local Kalman filter, and prevent divergence to improve robustness.

3.2. GMM for DKF fusion

The results of K Kalman filters are regarded as K Gaussian distributions, and a GMM model is used to represent the fusion of these distributions.

Consider a variable X with mean μ and covariance Σ following Gaussian mixture distribution. Based on Eq. (7), we get

$$\mu = \sum_{k=1}^{K} w_k \mu_k \tag{9}$$

$$\begin{aligned} \Sigma &= E(X^2) - \mu^2 \\ &= \textstyle\sum_{k=1}^K w_k(\mu_k^2 + \Sigma_k t) - \mu^2 \\ &= \textstyle\sum_{k=1}^K w_k \Sigma_k + [\sum_{k=1}^K w_k \mu_k^2 - \mu^2] \end{aligned} \tag{10}$$

Substitute the mean vector $\hat{x}^k_t$ and the covariance matrix $\hat{P}^k_t$ of the kth Kalman filter into Eqs. (9) and (10) to obtain

$$\hat{x}^m_t = \mu^m_t = \sum_{k=1}^{K} \hat{w}^k_t \hat{x}^k_t \tag{11}$$

$$\begin{aligned} \hat{P}^m_t &= \Sigma^m_t \\ &= \textstyle\sum_{k=1}^K \hat{w}^k_t \hat{P}^k_t + \left[\sum_{k=1}^K \hat{w}^k_t \hat{x}^k_t (\hat{x}^k_t)^T - \hat{x}^m_t \left(\hat{x}^m_t\right)^T\right] \end{aligned} \tag{12}$$

Based on EM algorithm [25], the weight w^k_t can be computed for the GMM composed of K Gaussian distributions with known means and covariances:

$$\gamma(z^{ik}_t) = \frac{w^k_t \mathcal{N}\left(x^i_t \mid \theta^k_t\right)}{\sum_{i=1}^K w^j_t \mathcal{N}\left(x^i_t \mid \theta^i_t\right)} \tag{13}$$

$$w^k_t = \frac{\sum_{i=1}^N \gamma(z^{ik}_t)}{N} \tag{14}$$

where N represents the number of samples from the kth Gaussian distribution, z_t^{ik} represents the latent variable, and $\theta_t^k = (\mu_t^k = \hat{x}_t^k, \Sigma_t^k = \hat{P}_t^k)$ are derived by local Kalman filter.

At timestep t, the weights of DKF can be computed by the closed-form solution. Based on Eq. (14), the weights are computed with known mean μ_k and covariance σ_k as

$$w_t^k = \frac{\sum_{j=1}^{K} \gamma_t^{jk}}{K} \tag{15}$$

Combining Eqs. (13) and (15), the weight can be expressed by the probability density function as

$$Kw_t^k = \sum_{j=1}^{K} \frac{w_t^k \mathcal{N}\left(x_t^j \mid \theta_t^k\right)}{\left(\sum_{i=1}^{K} w_t^i \mathcal{N}\left(x_t^j \mid \theta_t^i\right)\right)} \tag{16}$$

For each sensor node, there exists Eq. (16), which composes a multivariate equation set. The equation set can be decomposed of matrices in order to calculate the weights efficiently. Decompose the above equation into matrix form, and define five matrices:

$$W_t \triangleq [w_t^1, w_t^2, \ldots, w_t^K]_{K\times 1}^T \tag{17}$$

$$L_t \triangleq [L_t^1, L_t^2, \ldots, L_t^K]_{K\times 1}^T \tag{18}$$

$$\hat{K}_t \triangleq [K, K, \ldots, K]_{K\times 1}^T \tag{19}$$

$$S_t \triangleq \begin{bmatrix} \mathcal{N}\left(\hat{x}_t^1 \mid \hat{x}_t^1, \Sigma_t^1\right) & \cdots & \mathcal{N}\left(\hat{x}_t^K \mid \hat{x}_t^1, \Sigma_t^1\right) \\ \vdots & \ddots & \vdots \\ \mathcal{N}\left(\hat{x}_t^1 \mid \hat{x}_t^K, \Sigma_t^K\right) & \cdots & \mathcal{N}\left(\hat{x}_t^K \mid \hat{x}_t^K, \Sigma_t^K\right) \end{bmatrix}_{K\times K} \tag{20}$$

$$\hat{L}_t \triangleq [\frac{1}{L_t^1}, \frac{1}{L_t^2}, \ldots, \frac{1}{L_t^K}]_{K\times 1}^T \tag{21}$$

where

$$L_t^k \triangleq \frac{1}{\sum_{j=1}^{K} w_t^j \mathcal{N}(\hat{x}_t^k \mid \hat{x}_t^j, \Sigma_t^j)} \tag{22}$$

$\hat{L}_t$ contains the weighted components of the individual weight and distribution. W_t is the matrix of weights to be solved and $\hat{K}_t$ is extended to the same dimensions as W_t to facilitate the operation of the matrix. The column vector of S_t is composed of the numerator on the right-hand side of Eq. (16).

According to Eqs. (17)–(22), Eq. (16) can be re-written as

$$L_t = (S_t)^{-1}\hat{K}_t \tag{23}$$

$$\hat{L}_t = S_t^T W_t \tag{24}$$

The closed-form solution W_t can be calculated as

$$W_t = (S_t^T)^{-1}\hat{L}_t \tag{25}$$

Notice that Eq. (25) is the global optimal solution with $\sum_{k=1}^{K} w_k = 1$, and it has a unique solution. Since this is an over-constraint problem, the solution may not satisfy the condition $w_k \in [0, 1]$. When the global optimum cannot be reached, EM iteration with known mean and covariance can be utilized to obtain the local optimal solution.

3.3. The relationship between weight and covariance

In this subsection, we analyze the relationship between the weight and the covariance of LKF. The weight of distribution in GMM is solved by EM iteration with known mean and variance. Consider a GMM with one-dimension variable x_t, and the weights of all distributions are initialized to the same value w_0 at the first iteration:

$$w_0^1 = w_0^2 = \ldots = w_0^K = w_0 \tag{26}$$

$$w_1^k = \frac{w_0 \mathcal{N}(x_0^1 | x_0^k, {\sigma_0^k}^2)}{\sum_{i=1}^{K} w_0 \mathcal{N}(x_0^1 | x_0^i, {\sigma_0^i}^2)} \tag{27}$$

Let

$$D_1^k \triangleq \sum_{i=1}^{K} w_0 \mathcal{N}(x_0^1 | x_0^i, {\sigma_0^i}^2) \tag{28}$$

Then Eq. (27) can be expressed as:

$$w_1^k = \frac{w_0 \mathcal{N}(x_0^1 | x_0^k, {\sigma_0^k}^2)}{D_1^k} \tag{29}$$

At every timestep the local state and covariance are corrected by the fused results, so the estimation results of LKFs are close to the true value x_t^{true}. Assume $x_t^k = x_t^{true} + \delta_t^k$, where δ_t^k is sufficiently close to zero.

Since $D_1^1 = D_1^2 = \ldots = D_1^K$, expand the expression of the numerator based on Gaussian distribution probability density function as

$$w_0 \mathcal{N}(x_0^1 | x_0^k, {\sigma_0^k}^2) = \frac{w_0}{\sqrt{2\pi}} \times \frac{e^{-\frac{(\delta_0^k - \delta_0^1)^2}{2({\sigma_0^k}^2)}}}{\sigma_0^k} \tag{30}$$

The derivative of $w_0 \mathcal{N}(x_1 | x_k, \sigma_k^2)$ is

$$\frac{\partial w_0 \mathcal{N}(x_0^1 | x_0^k, {\sigma_0^k}^2)}{\partial \sigma_0^k} = \frac{w_0}{\sqrt{2\pi}} \times \frac{((\delta_0^k - \delta_0^1)^2 - {\sigma_0^k}^2) e^{-\frac{(\delta_0^k - \delta_0^1)^2}{2({\sigma_0^k}^2)}}}{{\sigma_0^k}^4} \tag{31}$$

Since $(\delta_0^k - \delta_0^1)^2$ is sufficiently close to zero, Eq. (31) is negative. Therefore, the w_1^k is negatively correlated with σ_0^k. In the jth iteration, the calculation of the weights can be expressed as

$$w_j^k = \frac{w_{j-1}^k \mathcal{N}(x_{j-1}^j | x_{j-1}^k, {\sigma_{j-1}^k}^2)}{\sum_{i=1}^{K} w_{j-1}^i \mathcal{N}(x_{j-1}^j | x_{j-1}^i, {\sigma_{j-1}^i}^2)} \tag{32}$$

$$D_j^k \triangleq \sum_{i=1}^{K} w_{j-1}^i \mathcal{N}(x_{j-1}^j | x_{j-1}^i, {\sigma_{j-1}^i}^2) \tag{33}$$

$$w_j^k = \frac{w_{j-1}^k \mathcal{N}(x_{j-1}^j | x_{j-1}^k, {\sigma_{j-1}^k}^2)}{D_j^k} \tag{34}$$

Based on Eq. (31), $\mathcal{N}(x_{j-1}^j | x_{j-1}^k, {\sigma_{j-1}^k}^2)$ is negatively correlated with σ_{j-1}^k. The weight of the distribution with large covariance will decrease after each iteration. For the multivariate model, the derivation is similar.

3.4. Index-Huber function

In order to dynamically reduce the influence of the large covariance, we design an index-Huber function to adjust the weights. The proposed index-Huber function is expressed as

$$\mathcal{H}(x) = \frac{x^\alpha}{\sqrt{x^{2\alpha-2} + \delta^{2\alpha-2}}}, x \in [0, 1] \tag{35}$$

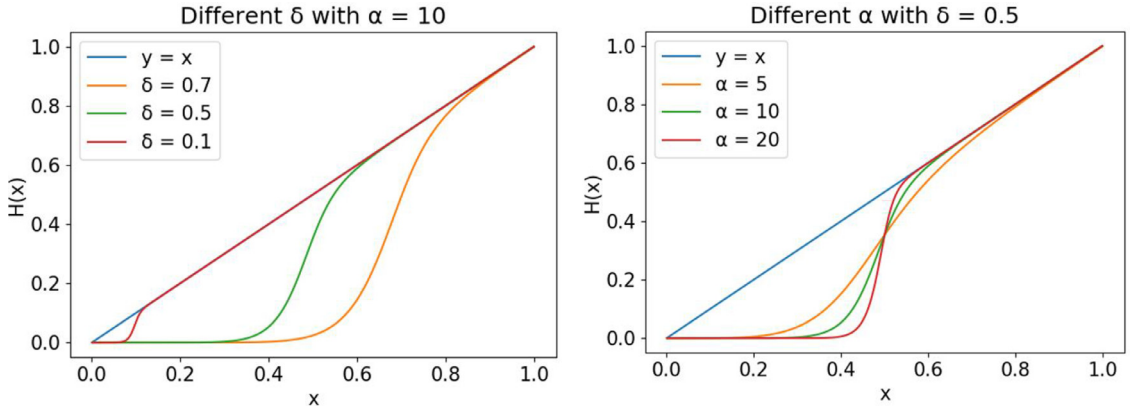


Fig. 2. Index-Huber function in Eq. (35) with different parameters. The properties of this function are mainly determined by the parameters δ and α.

where integer α ($\alpha \geq 2$) is the index parameter and $\delta \in (0, 1)$ is the minimal parameter. The index parameter α contributes to adjust the degree of weight scaling, and the minimal parameter δ is to set the threshold value. When $x \ll \delta$, Eq. (35) can be expressed as:

$$\mathcal{H}(x) = \frac{\frac{x^{\alpha}}{\delta^{\alpha-1}}}{\sqrt{\frac{x^{2\alpha-2}}{\delta^{2\alpha-2}} + 1}} \approx (\frac{x}{\delta})^{\alpha-1}x \tag{36}$$

When $x \gg \delta$, Eq. (35) can be expressed as:

$$\mathcal{H}(x) = \frac{x}{\sqrt{\frac{\delta^{2\alpha-2}}{x^{2\alpha-2}} + 1}} \approx x \tag{37}$$

The plots of $\mathcal{H}(x)$ with different α and δ are shown in Fig. 2. The parameter α affects the function slope in the interval $(0, \delta)$. According to the properties of $\mathcal{H}(x)$, it can be utilized to adjust the weights solved by EM algorithm, which can reduce the weight w_t^k when $w_t^k < \delta$ and maintain linearity when $w_t^k \gg \delta$.

3.5. Outlier detection

Since the outlier has an impact on weight calculation and fusion accuracy, the outlier detection is essential in the proposed framework. According to Eq. (31), when there existing an outlier in the distribution, the derivative of $w_0\mathcal{N}(x_1|x_k, \sigma_k^2)$ may be positive, which causes the weights to increase after iterations. The larger weight of the distribution with outliers can lead to the decrease in fusion accuracy, hence the outlier detection is necessary.

The outlier detection algorithms can be divided into two types: one is detecting the sensor measurement value, and the other is detecting the estimation results. In our framework, a measurement is detected as an outlier if the Mahalanobis distance exceeds a certain threshold [26].

$$y_{i,k}^T S_{i,k}^{-1} y_{i,k} > p \tag{38}$$

where $y_{i,k}$ is the transition function induced by the kinematic model, $S_{i,k}$ is the corresponding predicted covariance matrix and p is the threshold. If an outlier is detected, the distribution is rejected in this step. This method directly detects the outliers in sensor measurements with low computational cost. According to Eq. (38), the estimation results of LKF are guaranteed to be within the margin of error, which satisfied the condition that $x_t^k = x_t^{true} + \delta_t^k$. Different outlier detection algorithms can be used according to different sensor noises and sensor types. For example, the outlier detection of hydraulic servo drives in the presence of non-Gaussian noises is considered in [27].

3.6. Covariance intersection fusion

To improve the accuracy and reduce the cumulative error of LKF, batch covariance intersection (BCI) and sequential covariance intersection (SCI) are utilized in our framework. Compared with BCI algorithm which requires a massive computational cost, SCI can save computational cost. Although the accuracy of SCI is affected by the sensor order, in our framework only two covariance matrices are considered for fusion. The SCI algorithm in our method is expressed as:

$$\hat{P}_t^k = \left(\lambda_t^k \left(\hat{P}_{t|t}^k\right)^{-1} + \left(1 - \lambda_t^k\right)\left(\hat{P}_t^m\right)^{-1}\right)^{-1} \tag{39}$$

At timestep t, the covariance $\hat{P}_{t|t}^k$ of the kth LKF is fused with the covariance $\hat{P}_t^m$ of GMM. λ_t^k in Eq. (39) is an undetermined parameter which minimizes the performance index $J = \mathrm{tr}(\hat{P}_t^k)$.

$$\min J = \min_{\lambda_t^k \in [0,1]} \mathrm{tr}\,(\hat{P}_t^k) \tag{40}$$

Gold section method and Fibonacci method [28] can be used to find the optimal weighting coefficient λ_t^k [29]. Since λ_t^k minimizes $\mathrm{tr}(\hat{P}_t^k)$, we have

$$\mathrm{tr}(\hat{P}_t^k) \leq \mathrm{tr}(\hat{P}_t^m) \tag{41}$$

$$\mathrm{tr}(\hat{P}_t^k) \leq \mathrm{tr}(\hat{P}_{t|t}^k) \tag{42}$$

When the computational complexity is not considered, BCI can also be utilized to improve the accuracy of LKF. Fusing all covariance matrices of LKFs, the BCI is expressed as

$$\hat{P}_t^{BCI} = \left[\sum_{i=1}^{K} \lambda_t^i (\hat{P}_{t|t}^i)^{-1} + (1 - \sum_{i=1}^{K} \lambda_t^i)(\hat{P}_t^m)^{-1}\right]^{-1} \tag{43}$$

The undetermined parameters λ_t^i is to minimize the performance index $J_{BCI} = \mathrm{tr}(\hat{P}_t^{BCI})$.

$$\min J_{BCI} = \min_{\lambda_t^i \in [0,1]} \mathrm{tr}\,(\hat{P}_t^{BCI}) \tag{44}$$

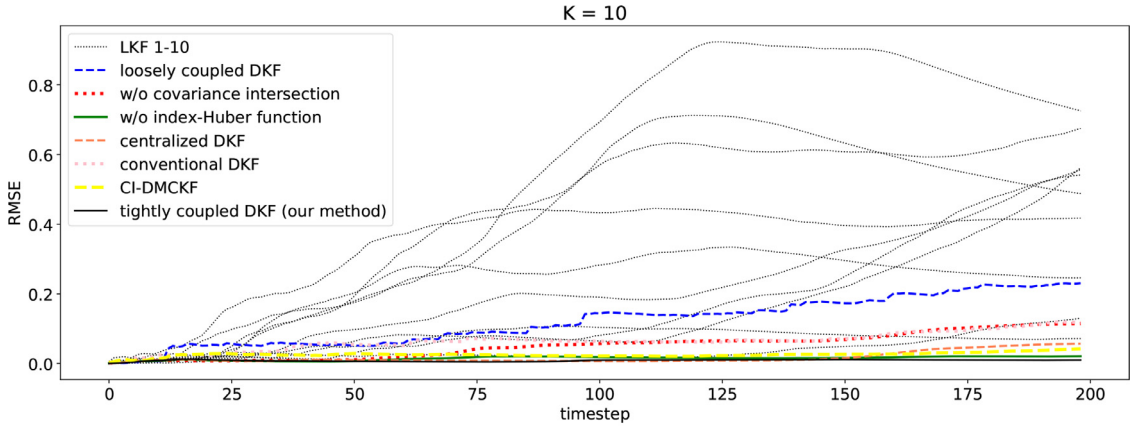
Solving for the parameters $\lambda_t^i (i = 1, 2, ., K)$ is a non-linear optimization problem that can be solved by sequential quadratic programming (SQP) algorithm [30].

4. Experiments

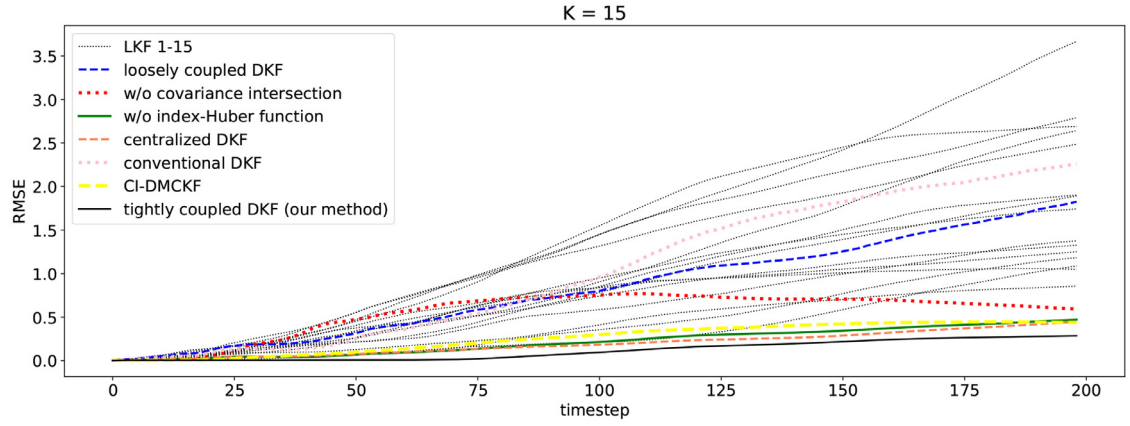
4.1. Simulation experiment

Consider a multi-sensor tracking system with K sensors:

$$x(t+1) = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} c(t) \tag{45}$$



(a) 10 sensor nodes



(b) 15 sensor nodes

Fig. 3. RMSE w.r.t. timestep of different methods in the simulation experiment with (a) 10 sensor nodes and (b) 15 sensor nodes. The experiment results demonstrate that our tightly coupled DKF outperforms other methods, and verify the effectiveness of the proposed covariance intersection and index-Huber function.

$$y_i(t) = Hx(t) + v_i(t), \quad i = 1, 2, \ldots, K \tag{46}$$

$$H = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{47}$$

In the simulation experiments, the sensors measure the velocity of the object with sampling interval $T_0 = 0.25s$. The noise $c(t)$ and $v_i(t)$ follow the non-Gaussian distribution, and the state x includes the velocity and the position. The initial parameters of local Kalman filter are set as: $R = \begin{bmatrix} 0.02 \end{bmatrix}$, $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $P_0 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. The Root Mean Squared Error (RMSE) reflects the deviation from the true value is utilized to evaluate the accuracy of the estimation results.

$$\text{RMSE} = \sqrt{\frac{1}{T_s}\sum_{t=1}^{T_s}(x_t - \hat{x}_t)^2} \tag{48}$$

where $\hat{x}_t$ is the true state. The RMSE results of different methods w.r.t. timestep are shown in Fig. 3, and the RMSE results at the final timestep are listed in Table 2, where tightly coupled DKF is the proposed method in this paper, conventional DKF, CI-DMCKF and centralized DKF are baseline methods in [7,11,31] respectively, loosely coupled DKF is the conventional DKF without centralized fusion, LKF 1–10 refers to the local sensor nodes, and w/o covariance intersection and w/o index-Huber function refer to tightly coupled DKF without covariance intersection and index-Huber function respectively. The experiment results demonstrate that the proposed method outperforms other methods, and the conventional DKF method [7] gets the worst performance, which does not consider non-Gaussian noises or nonlinearity. To capture the improvement in accuracy of the LKF by the tightly coupled structure, we calculate the cumulative error of the LKF to reflect the change in accuracy. As shown in Fig. 4, the cumulative error of loosely coupled LKF gradually increases, while the cumulative error of tightly coupled LKF does not significantly increase. This result demonstrates that tightly coupled structure can reduce the cumulative error of LKF.

Table 2
RMSE of Different Methods with 15 Sensor Nodes.

Method	Fused RMSE
Conventional DKF	2.2570
Centralized DKF	0.4449
DKF without covariance intersection (GMM)	0.5927
DKF without index-Huber function (GMM)	0.4711
Loosely coupled DKF (GMM)	1.8261
CI-DMCKF	0.4377
Tightly coupled DKF (GMM)	**0.2860**

The trace of the covariance matrix reflects the optimality of the estimation results [29]. Comparing the trace of the covariance matrix, the accuracy improvement of LKF by covariance intersection fusion can be evaluated. Table 3 shows the covariance matrix

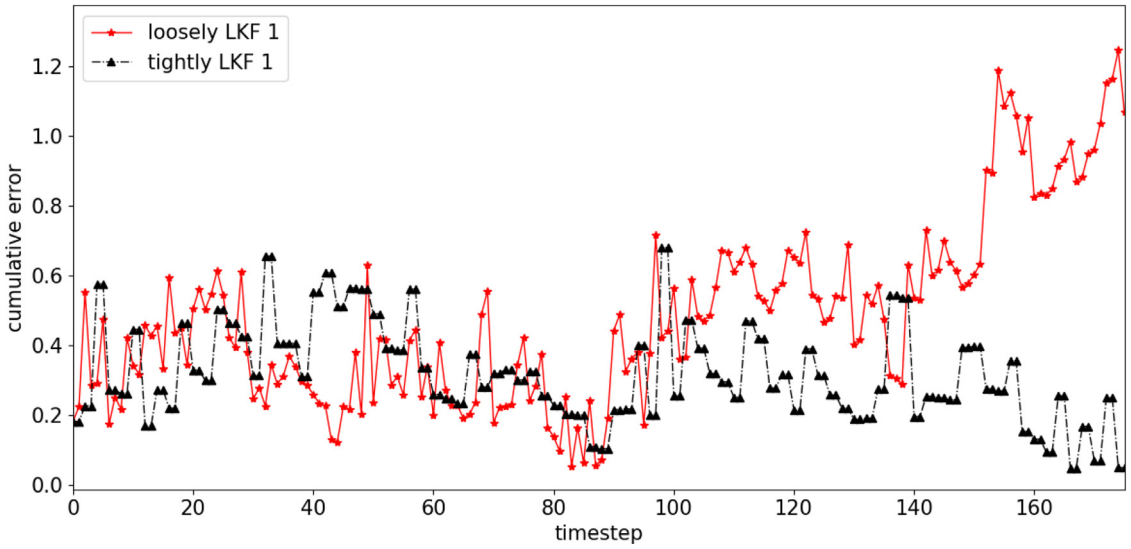


Fig. 4. Comparison of the cumulative error in the first LKF between tightly coupled DKF and loosely coupled DKF.

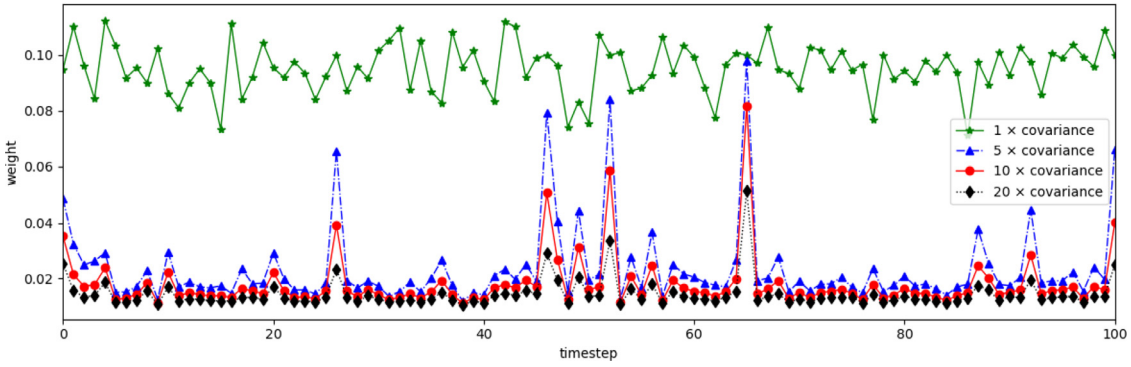


Fig. 5. The weight change with covariance increasing (Expand the covariance matrix of the first local Kalman filter by a multiple of 5, 10 and 20 respectively).

Table 3
The Trace of Covaiance Matrix.

T	GMM	LKF-1	C-LKF-1	LKF-10	C-LKF-10
20	9.691	7.475	**7.313**	12.553	**12.132**
40	16.723	14.540	**14.352**	20.330	**19.866**
60	24.646	22.208	**22.005**	28.256	**27.792**
80	33.699	30.318	**32.652**	36.473	**36.071**
100	41.773	38.697	**38.463**	44.951	**44.514**
120	50.910	47.319	**47.064**	53.197	**53.611**
140	60.192	55.939	**55.698**	62.264	**61.832**
160	68.906	64.674	**64.412**	70.999	**70.592**
180	77.410	73.381	**73.134**	79.734	**79.309**
200	87.160	82.227	**81.961**	88.575	**88.172**

trace of GMM and LKFs every 20 timesteps, where T refers to the timestep, C-LKF-1 refers to the covariance matrix trace of the first LKF with covariance intersection fusion, and LKF-10 refers to the covariance matrix trace of the 10th LKF without covariance intersection fusion. The experiments results demonstrate that the covariance intersection fusion reduces the trace of covariance matrix in the LKF. In other words, through covariance intersection fusion, the local estimation accuracy is improved.

In the above experiments, the covariance matrices of different LKFs are numerically close, so the effect of the index-Huber function on the weights is not obvious. Therefore, a controlled experiment is conducted which expands the covariance of an LKF. Fig. 5 shows the weight change of the corresponding distribution as the covariance increases. When the covariance increases to five times, the weight of the distribution decreases from 0.103 to 0.028. As shown in Fig. 6, the index-Huber function improves the estimation accuracy of tightly coupled DKF algorithm, and with the increase of the expansion factor, the index-Huber function has a more significant improvement in accuracy.

4.2. Experiment on a real-world system

Tightly coupled DKF method is verified on a complex hexapod robot system. This system, as shown in Fig. 7, contains 3 IMU sensors, 18 torque sensors, and 18 joint encoders. Due to the quaternion differentiation, the kinematics differential equation of this system has strong nonlinearity. Since the IMU sensor measurement will have high energy noise when the robot is walking, the noises follow the non-Gaussian distribution [32].

To estimate the body position r, velocity v, orientation q, and the positions of 6 foot contact points p_i for the robot, two local extended Kalman filters which have different kinematics assumptions are established. In the first extended Kalman filer (EKF), assume that the robot is always in uniform acceleration motion, which is similar to [26]. The discrete-time equations can be expressed as

$$\hat{r}_{k+1} = \hat{r}_k + \Delta t \hat{v}_k + \frac{\Delta t^2}{2}\left(\hat{C}_k^T \hat{f}_k + g\right) \tag{49}$$

$$\hat{v}_{k+1} = \hat{v}_k \Delta t\left(\hat{C}_k^T \hat{f}_k + g\right) \tag{50}$$

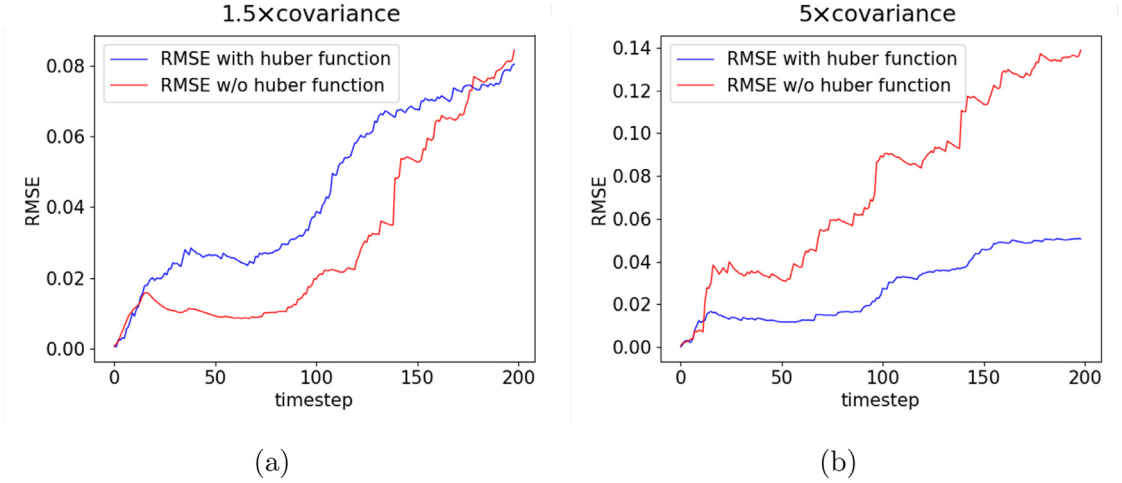


Fig. 6. The effect of index-Huber function on RMSE in different cases: (a) Covariance expansion by a factor of 1.5, and (b) by a factor of 5.

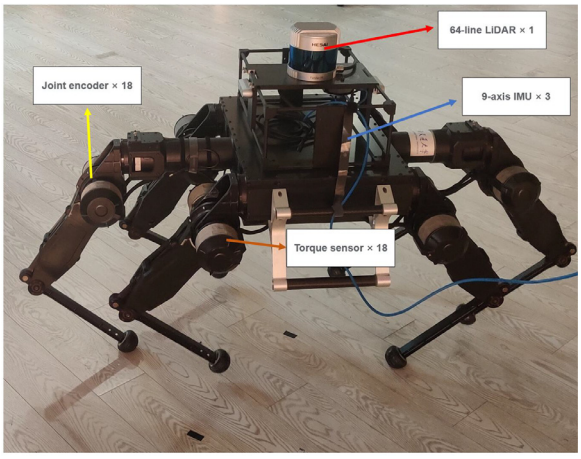


Fig. 7. The real-world robot system with multiple sensors, including 3 IMU sensors, 18 torque sensors, and 18 joint encoders.

$$\hat{q}_{k+1} = \zeta\left(\Delta t \hat{\omega}_k\right) \otimes \hat{q}_k \tag{51}$$

$$\hat{p}_{i,k+1} = \hat{p}_{i,k} \quad \forall i \in \{1, \ldots, 6\} \tag{52}$$

where $\hat{C}_k^T$ represents the transformation which defines the inertial coordinate frame I relative to the body coordinate frame B; The IMU measurements $\hat{f}_k$ and $\hat{\omega}_k$ refer to the acceleration and angular velocity respectively. Equation (51) is the incremental rotation equation of quaternion. In the second EKF, assume that the robot is always moving at a constant speed, which is similar to Yang et al. [32]. The discrete-time equations are:

$$\hat{r}_{k+1} = \hat{r}_k + \Delta t \hat{v}_k \tag{53}$$

$$\hat{v}_{k+1} = \hat{v}_k \tag{54}$$

$$\hat{q}_{k+1} = \hat{q}_k + \frac{1}{2} q_t \otimes \begin{bmatrix} \left(\hat{\omega}_k\right) \Delta t \\ 0 \end{bmatrix} \tag{55}$$

The velocity $\hat{v}_k$ and angular velocity $\hat{\omega}_k$ in this method are computed through the joint encoders measurements, and the calculation process is same with [32]. The specific Kalman filter equations of two methods can obtained in [26] and [32] respectively.

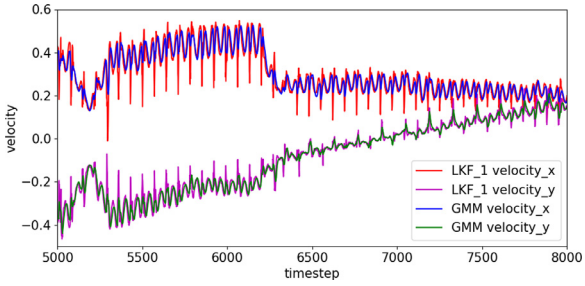


Fig. 8. The estimated velocity of the robot from T-DKF (tightly coupled DKF) and L-DKF (loosely coupled DKF) in different timesteps. The velocity estimated directly from the IMU fluctuates considerably, but the GMM fusion results reduce the effect of IMU noise.

Table 4
RMSE of Different Methods at Different Timesteps.

T	LKF 1	LKF 2	L-DKF	T-DKF
4000	0.6226	0.6188	0.5272	**0.4831**
6000	0.9989	0.9906	0.7961	**0.6654**
8000	1.2004	1.1898	0.8023	**0.7094**
10,000	1.2088	1.1978	0.7847	**0.6966**
12,000	1.1291	1.1186	0.7892	**0.6982**
14,000	1.0439	1.0339	0.8167	**0.7348**
16,000	0.9627	0.9534	0.8372	**0.7519**
18,000	0.8662	0.8576	0.8239	**0.7495**

The true state of the robot movement is provided by a lidar odometry [33]. The sampling frequency in this experiment is 200Hz, and Fig. 8 shows the estimated velocity in different timesteps. The IMU measurements have high-energy noise when robot contacts the ground, which causes the velocity measurements to be accompanied by high-frequency noises. The GMM-based fusion suppresses the influence of the high frequency noise from IMU measurements. Comparisons of RMSE in different timesteps are shown in Table 4. The RMSE of GMM is always the smallest, and the accuracy of the local Kalman filter is also improved after covariance intersection.

5. Conclusion

In this paper, we propose a tightly coupled DKF fusion algorithm based on covariance intersection and GMM. The Monte-Carlo simulation experiments show that the proposed method can improve the accuracy of estimation results and outperforms other methods. The simulation experiments show that the tightly cou-

pled method can reduce the cumulative error of LKF to improve the overall accuracy, and the designed index-Huber function can mitigate the influence caused by large covariance. A real-world hexapod robot with multiple sensors is used to evaluate the proposed method. The experiment demonstrates that the proposed method can improve the fusion accuracy on a complex system. In future works, the influence of covariance on the fusion results should also be considered when solving the GMM weights.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yangqing Fu: Data curation, Methodology, Writing – original draft, Software. **Ming Sun:** Data curation, Writing – review & editing. **Yue Gao:** Supervision, Methodology, Writing – review & editing.

References

[1] H. Zhang, X. Zhou, Z. Wang, H. Yan, Maneuvering target tracking with event-based mixture Kalman filter in mobile sensor networks, IEEE Trans. Cybern. 50 (10) (2019) 4346–4357.

[2] W. Li, Y. Jia, Distributed consensus filtering for discrete-time nonlinear systems with non-Gaussian noise, Signal Process. 92 (10) (2012) 2464–2470.

[3] S. Xu, K. Doğançay, H. Hmam, Distributed pseudolinear estimation and UAV path optimization for 3D AOA target tracking, Signal Process. 133 (2017) 64–78.

[4] S. Sun, H. Lin, J. Ma, X. Li, Multi-sensor distributed fusion estimation with applications in networked systems: a review paper, Inf. Fusion 38 (2017) 122–134.

[5] H. Lin, C. Hu, Z. Deng, G. Liu, Distributed Kalman filter with fuzzy noises over multi-agent systems, IEEE Trans. Fuzzy Syst. (2021).

[6] M. Hassan, G. Salut, M. Singh, A. Titli, A decentralized computational algorithm for the global Kalman filter, IEEE Trans. Automat. Control 23 (2) (1978) 262–268.

[7] J. Sijs, M. Lazar, P. Van den Bosch, Z. Papp, An overview of non-centralized Kalman filters, in: 2008 IEEE International Conference on Control Applications, IEEE, 2008, pp. 739–744.

[8] S.B. Kanagala, K.P. Detroja, Distributed state estimation through co-acting Kalman filters, Asian J. Control 23 (5) (2021) 2495–2506.

[9] A. Ahmad, M. Gani, F. Yang, Decentralized robust Kalman filtering for uncertain stochastic systems over heterogeneous sensor networks, Signal Process. 88 (8) (2008) 1919–1928.

[10] H. Long, Z. Qu, X. Fan, S. Liu, Distributed extended Kalman filter based on consensus filter for wireless sensor network, in: Proceedings of the 10th World Congress on Intelligent Control and Automation, IEEE, 2012, pp. 4315–4319.

[11] S.P. Talebi, S. Werner, Distributed Kalman filtering and control through embedded average consensus information fusion, IEEE Trans. Automat. Control 64 (10) (2019) 4396–4403.

[12] W. Li, G. Wei, F. Han, Y. Liu, Weighted average consensus-based unscented Kalman filtering, IEEE Trans. Cybern. 46 (2) (2015) 558–567.

[13] W. Song, Z. Wang, J. Wang, J. Shan, Particle filtering for a class of cyber-physical systems under round-robin protocol subject to randomly occurring deception attacks, Inf. Sci. 544 (2021) 298–307.

[14] V. Stojanović, D. Pršić, L. Dubonjić, Joint estimation of states and parameters of linear systems with parameter faults under non-Gaussian noises, Facta Univ. Ser. Autom. ControlRob. 18 (2) (2020) 113–125.

[15] J. Chen, J. Li, S. Yang, F. Deng, Weighted optimization-based distributed Kalman filter for nonlinear target tracking in collaborative sensor networks, IEEE Trans. Cybern. 47 (11) (2016) 3892–3905.

[16] X. Sheng, Y.-H. Hu, P. Ramanathan, Distributed particle filter with GMM approximation for multiple targets localization and tracking in wireless sensor network, in: IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005., IEEE, 2005, pp. 181–188.

[17] G.J. McLachlan, S. Rathnayake, On the number of components in a Gaussian mixture model, Wiley Interdiscip. Rev. Data Min.Knowl. Discov. 4 (5) (2014) 341–355.

[18] R. Singh, B.C. Pal, R.A. Jabr, Statistical representation of distribution system loads using Gaussian mixture model, IEEE Trans. Power Syst. 25 (1) (2009) 29–37.

[19] W. Li, Y. Jia, J. Du, F. Yu, Gaussian mixture PHD filter for multi-sensor multi–target tracking with registration errors, Signal Process. 93 (1) (2013) 86–99.

[20] N.Y. Ko, T.G. Kim, Comparison of Kalman filter and particle filter used for localization of an underwater vehicle, in: 2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), IEEE, 2012, pp. 350–352.

[21] R. Olfati-Saber, Distributed kalman filtering for sensor networks, in: 2007 46th IEEE Conference on Decision and Control, IEEE, 2007, pp. 5492–5498.

[22] G. Wang, N. Li, Y. Zhang, Diffusion distributed Kalman filter over sensor networks without exchanging raw measurements, Signal Process. 132 (2017) 1–7.

[23] S.J. Julier, J.K. Uhlmann, New extension of the Kalman filter to nonlinear systems, in: Signal Processing, Sensor Fusion, and Target Recognition VI, vol. 3068, International Society for Optics and Photonics, 1997, pp. 182–193.

[24] D.A. Reynolds, Gaussian mixture models, Encycl. Biom. 741 (2009) 659–663.

[25] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the em algorithm, J. R. Stat. Soc. Series B (Methodological) 39 (1) (1977) 1–22.

[26] M. Bloesch, C. Gehring, P. Fankhauser, M. Hutter, M.A. Hoepflinger, R. Siegwart, State estimation for legged robots on unstable and slippery terrain, in: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2013, pp. 6058–6064.

[27] V. Stojanovic, D. Prsic, Robust identification for fault detection in the presence of non-Gaussian noises: application to hydraulic servo drives, Nonlinear Dyn. 100 (3) (2020) 2299–2313.

[28] W. Sun, Y.-X. Yuan, Optimization Theory and Methods: Nonlinear Programming, Vol. 1, Springer Science & Business Media, 2006.

[29] Z. Deng, P. Zhang, W. Qi, J. Liu, Y. Gao, Sequential covariance intersection fusion Kalman filter, Inf. Sci. 189 (2012) 293–309.

[30] A. Ruszczynski, Nonlinear optimization, Nonlinear Optimization, Princeton University Press, 2011.

[31] C. Hu, B. Chen, An efficient distributed Kalman filter over sensor networks with maximum correntropy criterion, IEEE Trans. Signal Inf. Process. Netw. (2022).

[32] S. Yang, H. Kumar, Z. Gu, X. Zhang, M. Travers, H. Choset, State estimation for legged robots using contact-centric leg odometry, arXiv:1911.05176 (2019).

[33] T. Shan, B. Englot, LeGO-LOAM: lightweight and ground-optimized lidar odometry and mapping on variable terrain, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2018, pp. 4758–4765.