Received December 10, 2019, accepted January 16, 2020, date of publication January 22, 2020, date of current version January 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968602

Distributed Robust Cubature Information Filtering for Measurement Outliers in Wireless Sensor Networks

JIAHAO ZHANG [1], SHESHENG GAO [1], XIAOMIN QI [2], JIAHUI YANG [1], JUAN XIA [1], AND BINGBING GAO [1]

[1]Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China
[2]Department of Electrical and Computer Engineering, COMSATS University Islamabad, Abbottabad Campus, Abbottabad 22060, Pakistan

Corresponding author: Jiahao Zhang (zjh_dr@126.com)

This work was supported in part by the National Nature Science Foundation of China under Grant 41904028, in part by the Science, Technology and Innovation Commission of Shenzhen Municipality, China, under Grant JCYJ20180306171439979, and in part by the Shaanxi Province Key Research and Development Projects, China, under Grant 2018ZDXM-GY-024.

ABSTRACT In wireless sensor networks (WSN), measurements are always corrupted by outliers or impulsive noise. Cubature information filtering (CIF) is founded based on minimum mean square error (MMSE) criterion, which is not applicable to non-Gaussian noise. Hence, a novel robust CIF (RCIF) is derived based on maximum correntropy criterion (MCC) to enhance the robustness of state estimation in the local node. For the information fusion, weighted average consensus (WAC) based distributed RCIF (DRCIF) is founded to improve the stability of sensor networks and the accuracy of state estimation. The estimation error of DRCIF is proved to be bounded in mean square. Numerical simulations are provided to evaluate the effectiveness of proposed algorithms.

INDEX TERMS Robust cubature information filtering, maximum correntropy criterion, Non-Gaussian measurement noise, distributed state estimation, weighted average consensus.

I. INTRODUCTION

With the development of communication, cloud computing and embedded technology, wireless sensor technology has been getting increased attention in recent years [1], and WSN is gradually applied to navigation, environment monitoring, and target tracking etc. Based on the difference of information fusion, WSN can be divided into three groups [2], i.e., centralized, decentralized and distributed WSN. Normally, data of all sensor nodes must be transmitted into the centralized computing node to obtain optimal results of information fusion in centralized WSN [3]. Nevertheless, limited by the bandwidth of WSN and the computing capacity of the centralized evaluate node, centralized WSN is usually not feasible in practice, especially for the large scale networks. The mode of single compute node is abandoned by the decentralized WSN and all sensor nodes in networks are regarded as the evaluate node [4]. Although the robustness of networks is enhanced, all senor nodes in decentralized WSN are forced to keep in touch with each other and the limited network bandwidth is still an unsolved problem. Cooperative mechanism of neighbor nodes is employed in distributed WSN to reduce the need of substantial network bandwidth and strong evaluate capacity in the single node, and then the robustness and stability of networks are enhanced consequently. Hence, the distributed architecture WSN gradually becomes the mainstream in using [5].

Distributed state estimation in WSN was investigated in [6]–[15]. Distributed Kalman filtering (DKF) was derived in [6], [7] to obtain the optimal estimation. However, DKF is only applicable to the linear system, and the system model is nonlinear in real applications. On the basis of DKF, distributed extended Kalman filtering (DEKF) was designed for the nonlinear system [8], [9]. Nevertheless, large estimation errors may exist in DEKF due to the first-order linearization of nonlinear system models [16]. Unscented transformation was adopted in distributed unscented Kalman filtering (DUKF) [10], [11] to acquire more accurate estimations than DEKF. But the covariance matrix of DUKF may be non-positive in high-dimensional system [17]. To overcome the drawback of DUKF, distributed cubature Kalman filtering (DCKF) [12], [13] and distributed cubature information

The associate editor coordinating the review of this manuscript and approving it for publication was Feiqi Deng.

filtering (DCIF) [14], [15] were founded based on spherical cubature integration rule. DCIF is the algebraic equivalent form of DCKF, and both are designed for state estimation in nonlinear systems. DCIF can provide more stable and accurate estimations than most Gaussian filters [18], such as DEKF and DUKF. Moreover, compared with Kalman filtering, information filtering is more suitable for distributed information fusion in WSN [19]. In general, the above methods for distributed state estimation falls into three types [13]: consensus on estimates (CE), consensus on measurements (CM) and consensus on information (CI). [9] belongs to the scope of CE, which only focuses on the information fusion of state estimation in local nodes, but regardless of the fusion of error covariance [20]. Hence, the filtering performance of CE based methods usually cannot be enhanced. The information fusion of local measurements and innovation covariance is included in the CM based methods [6], [7], [10], [12], but the stability can only be guaranteed by abundant consensus steps [20]. CI is derived based on information filtering and WAC based methods [11], [14], [15] belong to the scope of CI. The procedure of information fusion in CI includes the fusion of information state vector and information matrix, and the stability of CI is not affected by the times of consensus steps [14]. The information fusion in [8], [13] is constructed by combining CM and CI to obtain more accurate estimations, but the calculation is increased with the incremental communication bandwidth. Hence, in terms of stability and precision, CI is more suitable for information fusion in distributed WSN.

In WSN, outliers or impulse noise that caused by electromagnetic interference or communication failure usually cannot be eliminated among measurements. This situation always results in large errors in estimations or even divergence of state estimator, especially for extremely large measurement outliers. The above methods for distributed state estimation are founded based on MMSE criterion, which cannot cope with state evaluation under non-Gaussian noise. In recent years, MCC based Kalman filtering [21], [22] is derived to deal with non-Gaussian noise for linear systems. Correntropy which is sensitive to pulse can be used to survey the local similarity of measurements. Hence, MCC can be selected as the cost equation to design corresponding filtering methods for abnormal measurements. For nonlinear systems, MCC based unscented Kalman filtering and cubature Kalman filtering were investigated in [23], [24] and [25], respectively. However, the designed MCC based nonlinear methods is only suitable for state estimation in a single sensor node, but not applicable for state evaluation in distributed WSN. A distributed cubature information filtering based on MCC is designed in [26] to cope with measurement outliers in WSN. Nevertheless, the DCIF cannot deal with state estimation under extremely large measurement outliers, and estimation errors of the distributed CIF are not proved to be bounded in mean square.

In this work, we present a novel method for state estimation under measurement outliers or impulse noise in WSN. RCIF is designed based on MCC to cope with measurement outliers or impulse noise in the local sensor node. Then, WAC based DRCIF is derived for distributed information fusion in WSN to enhance the accuracy of state estimation and the stability of WSN. The main contributions of this paper are given as follows:

1) A novel distributed robust filtering is derived for nonlinear systems to cope with measurement outliers in WSN, particularly extremely large outliers.

2) The stochastic boundedness of estimation errors of DRCIF is investigated.

The structure of this paper is given as follows. System model is presented in section 2. Section 3 provides the derivation of RCIF. DRCIF is designed in section 4. Estimation errors of DRCIF are proved to be bounded in section 5. Results of different simulations are discussed in section 6 to evaluate the performance of the proposed method. Conclusions are drawn in section 7.

II. SYSTEM MODEL

In this section, nonlinear discrete time system is chosen to model distributed WSN. The process equation and measurement equation are displayed by the first and second term in (1), respectively.

$$\begin{cases} \mathbf{X}_k = f(\mathbf{X}_{k-1}) + \omega_{k-1} \\ \mathbf{Z}_k^n = h^n(\mathbf{X}_k) + \upsilon_k^n, \quad n = 1, \cdots, N \end{cases} \tag{1}$$

where $\mathbf{X}_k \in \mathbb{R}^m$ is the state vector at discrete-time instant k. $N \geqslant 2$ is the number of sensor nodes, $\mathbf{Z}_k^n \in \mathbb{R}^l$ is the measurement vector of node n at discrete-time instant k. $f(\cdot)$ denotes the nonlinear state transition function. $h^n(\cdot)$ denotes the nonlinear measurement function of node n. $\omega_k \in \mathbb{R}^m$ denotes the process noise and $\upsilon_k^n \in \mathbb{R}^l$ denotes the measurement noise, and they are assumed as uncorrelated Gaussian noise, namely $\omega_k \sim N(0, \mathbf{Q}_k)$ and $\upsilon_k^n \sim N(0, \mathbf{R}_k)$, where $C \sim N(0, \mathbf{D})$ denotes that C is zero-mean Gaussian white noise with covariance matrix $\mathbf{D}$. The distributed WSN adopts cooperative mechanism of adjacent nodes. Hence, the undirected graph of distributed WSN is presented as $\Psi(\mathbb{N}, \varepsilon)$, where $\mathbb{N} = \{1, \cdots, N\}$ denotes the node set. If node j can accept the data transmitted from node n, then $(n, j) \in \varepsilon$. For node n, $\mathbb{N}_n = \{n | (n, j) \in \varepsilon\}$ denotes the adjacent nodes set. If there are no adjacent nodes of node n, then $\mathbb{N}_n = \emptyset$.

III. ROBUST CIF

This section mainly focuses on the derivation of RCIF, including the inference of cost function and the algorithm flow.

A. COST FUNCTION OF RCIF

For random variables A and B, correntropy is defined by

$$V(A, B) = E_{A,B}[\kappa(A, B)] = \int\int \kappa(A, B)\, dF_{A,B}(A, B) \tag{2}$$

where $E[\centerdot]$ denotes expectation. $\kappa(\centerdot)$ denotes kernel function. $F_{A,B}(A, B)$ denotes the joint probability distribution of variables A and B. The expression of correntropy in discrete time is approximated by

$$\hat{V}(A,B)=E_{A,B}[\kappa(A,B)]\approx\frac{1}{m}\sum_{i=1}^{m}\kappa(A_i,B_i) \tag{3}$$

where m denotes the dimension of variables A and B. Gaussian kernel is usually selected as kernel function

$$\kappa_\sigma(e_i)=\exp(-\frac{e_i^2}{2\sigma^2}) \tag{4}$$

where $e_i=A_i-B_i$, and σ denotes the bandwidth of Gaussian kernel. $\kappa_\sigma(e_i)$ reaches the maxima when $e_i=0$. Then the cost function of MCC is defined by

$$C_{MCC}=\min\left(\sum_{i=1}^{m}(\kappa_\sigma(0)-\kappa_\sigma(e_i))\right) \tag{5}$$

To simplify the derivation of cost function for RCIF, pseudo-measurement matrix is constructed by [27]

$$\mathbf{H}_k^n=\left(\mathbf{P}_{xz,k|k-1}^n\right)^T\left(\mathbf{P}_{k|k-1}\right)^{-1},\quad n=1,\cdots,N \tag{6}$$

where $\mathbf{P}_{k|k-1}$ is the error covariance matrix of predicted state. $\mathbf{P}_{xz,k|k-1}^n$ is the cross-covariance matrix of predicted state and predicted measurement. Meanwhile, to compensate the linearization error of measurement equation, diagonal matrix $\mu_k^n=diag\left[\mu_{k,1}^n\ \mu_{k,2}^n\ \cdots\ \mu_{k,l}^n\right]$ is introduced. The linearization of measurement equation in (1) is defined by

$$\mathbf{Z}_k^n=\mu_k^n\mathbf{H}_k^n\mathbf{X}_k+\upsilon_k^n,\quad n=1,\cdots,N \tag{7}$$

where $h^n(\mathbf{X}_k)=\mu_k^n\mathbf{H}_k^n\mathbf{X}_k$. The state estimation of Kalman based filtering methods is usually equivalent to search the following minimization problem [28]

$$\hat{\mathbf{X}}_k=\arg\min\left(\left\|\mathbf{X}_k-\hat{\mathbf{X}}_{k|k-1}\right\|^2_{\mathbf{P}_{k|k-1}^{-1}}+\left\|\mathbf{Z}_k^n-\hat{\mathbf{Z}}_k^n\right\|^2_{(\mathbf{R}_k^n)^{-1}}\right) \tag{8}$$

where $\hat{\mathbf{X}}_{k|k-1}$ and $\hat{\mathbf{Z}}_k$ denote the predicted state and measurement estimation at discrete-time instant k, respectively. $\mathbf{R}_k^n$ is the error covariance matrix of measurement noise. For the nonlinear measurement model, linearization error of measurement equation must be compensated to improve precision of state estimation and $\hat{\mathbf{Z}}_k^n=\mu_k^n\mathbf{H}_k^n\mathbf{X}_k$; for the linear measurement model, $\hat{\mathbf{Z}}_k^n=\mathbf{H}_k^n\mathbf{X}_k$ and $\mu_k^n=\mathbf{I}^{l\times l}$. Denote the residual error item

$$\zeta_k^n=(\mathbf{R}_k^n)^{-1/2}(\mu_k^n\mathbf{H}_k^n\mathbf{X}_k-\mathbf{Z}_k^n) \tag{9}$$

where $\zeta_k^n=\left[\zeta_{k,1}^n,\ \ldots,\ \zeta_{k,l}^n\right]^T$, l is the dimension of measurement vectors. Cost function in (5) is used to redefine the minimization problem

$$\hat{\mathbf{X}}_k=\arg\min\left(\left\|\mathbf{X}_k-\hat{\mathbf{X}}_{k|k-1}\right\|^2_{\mathbf{P}_{k|k-1}^{-1}}+\sum_{i=1}^{l}\left(\kappa_\sigma(0)-\kappa_\sigma(\zeta_{k,i}^n)\right)\right) \tag{10}$$

Take the partial derivative of (10) with respect to $\mathbf{X}_k$, one has that

$$\mathbf{P}_{k|k-1}^{-1}\left(\mathbf{X}_k-\hat{\mathbf{X}}_{k|k-1}\right)-\sum_{i=1}^{l}\left(\frac{\partial\kappa_\sigma\left(\zeta_{k,i}^n\right)}{\partial\zeta_{k,i}^n}\frac{\partial\zeta_{k,i}^n}{\partial\mathbf{X}_k}\right)=0 \tag{11}$$

Substitute (4) into (11)

$$\left(\mathbf{P}_{k|k-1}\right)^{-1}\left(\mathbf{X}_k-\hat{\mathbf{X}}_{k|k-1}\right)+\frac{1}{\sigma^2}\sum_{i=1}^{D}\exp\left(-\frac{\left(\zeta_{k,i}^n\right)^2}{2\sigma^2}\right)\zeta_{k,i}^n\frac{\partial\zeta_{k,i}^n}{\partial\mathbf{X}_k}=0 \tag{12}$$

Denote

$$\psi_i=\frac{1}{\sigma^2}\exp\left(-\frac{\left(\zeta_{k,i}^n\right)^2}{2\sigma^2}\right),\quad i=1,\cdots,l \tag{13}$$

$$\psi=diag[\psi_1\ \cdots\ \psi_l] \tag{14}$$

The partial derivative of ζ_k^n is defined by

$$\frac{\partial\zeta_k^n}{\partial\mathbf{X}_k}=(\mathbf{R}_k^n)^{-1/2}(\mu_k^n\mathbf{H}_k^n) \tag{15}$$

Substitute (15) and ζ_k^n into (12), one has that

$$\left(\mathbf{P}_{k|k-1}\right)^{-1}\left(\mathbf{X}_k-\hat{\mathbf{X}}_{k|k-1}\right)+(\mu_k^n\mathbf{H}_k^n)^T(\mathbf{R}_k^n)^{-T/2}\psi(\mathbf{R}_k^n)^{-1/2}(\mu_k^n\mathbf{H}_k^n\mathbf{X}_k-\mathbf{Z}_k^n)=0 \tag{16}$$

Eq. (16) satisfies the minimization solution for the cost function of Kalman based nonlinear filtering, like RCIF, and the minimization solution is redefined by

$$\hat{\mathbf{X}}_k=\arg\min\left(\left\|\mathbf{X}_k-\hat{\mathbf{X}}_{k|k-1}\right\|^2_{\mathbf{P}_{k|k-1}^{-1}}+\left\|\mathbf{Z}_k^n-\mu_k^n\mathbf{H}_k^n\mathbf{X}_k\right\|^2_{(\bar{\mathbf{R}}_k^n)^{-1}}\right) \tag{17}$$

where $\bar{\mathbf{R}}_k^n$ is the adjusted covariance matrix of measurement noise and defined by

$$\bar{\mathbf{R}}_k^n=(\mathbf{R}_k^n)^{1/2}\psi^{-1}(\mathbf{R}_k^n)^{T/2} \tag{18}$$

Since $h^n(\mathbf{X}_k)=\mu_k^n\mathbf{H}_k^n\mathbf{X}_k$, the equivalent form of (17) is written as

$$\hat{\mathbf{X}}_k=\arg\min\left(\left\|\mathbf{X}_k-\hat{\mathbf{X}}_{k|k-1}\right\|^2_{\mathbf{P}_{k|k-1}^{-1}}+\left\|\mathbf{Z}_k^n-h^n(\mathbf{X}_k)\right\|^2_{\bar{\mathbf{R}}_k^{-1}}\right) \tag{19}$$

B. THE GENERAL FORM OF RCIF

For the local node n, the general form of RCIF is summarized as follows.

1) TIME UPDATE

For node n, the state estimation $\hat{\mathbf{X}}_{k-1}^n$ and information matrix $\mathbf{Y}_{k-1}^n$ are supposed known at time instant $k-1$. According to the spherical cubature integration rule, sigma points are generated by

$$\chi_{k-1}^{n,i} = \hat{\mathbf{X}}_{k-1}^n + \sqrt{\left(\mathbf{Y}_{k-1}^n\right)^{-1}}\xi^i, \quad i = 1, \cdots, 2m \tag{20}$$

where m is the dimension of state vector $\hat{\mathbf{X}}_k^n$. $\sqrt{\left(\mathbf{Y}_{k-1}^n\right)^{-1}}$ is a square root matrix defined by $\left(\mathbf{Y}_{k-1}^n\right)^{-1} = \left(\mathbf{Y}_{k-1}^n\right)^{-1/2}\left(\mathbf{Y}_{k-1}^n\right)^{-T/2}$.
$\xi^i = \begin{cases} \sqrt{m}\mathbf{e}_i, & i = 1, \cdots, m \\ -\sqrt{m}\mathbf{e}_i, & i = n+1, \cdots, 2m \end{cases}$, where $\mathbf{e}_i$ denotes m-dimensional unit vector with *ith* element being 1.

Then, $\chi_{k-1}^{n,i}$ is mapped by the nonlinear state transition function

$$\chi_{k|k-1}^{n,i} = f\left(\chi_{k-1}^{n,i}\right), \quad i = 1, \cdots, 2m \tag{21}$$

Estimate the predicted state $\hat{\mathbf{X}}_{k|k-1}^n$ and its error covariance matrix $\mathbf{P}_{k|k-1}^n$, respectively, as

$$\hat{\mathbf{X}}_{k|k-1}^n = \frac{1}{2m}\sum_{i=1}^{2m}\chi_{k|k-1}^{n,i} \tag{22}$$

$$\mathbf{P}_{k|k-1}^n = \frac{1}{2m}\sum_{i=1}^{2m}(\chi_{k|k-1}^{n,i} - \hat{\mathbf{X}}_{k|k-1}^n)(\chi_{k|k-1}^{n,i} - \hat{\mathbf{X}}_{k|k-1}^n)^T + \mathbf{Q}_{k-1} \tag{23}$$

Next, the predicted information matrix and predicted information state vector are written by

$$\mathbf{Y}_{k|k-1}^n = \left(\mathbf{P}_{k|k-1}^n\right)^{-1} \tag{24}$$

$$\hat{\mathbf{y}}_{k|k-1}^n = \mathbf{Y}_{k|k-1}^n\hat{\mathbf{X}}_{k|k-1}^n \tag{25}$$

2) MEASUREMENT UPDATE

Based on (22) and (23), a new set of sigma points is generated by

$$\bar{\chi}_{k|k-1}^{n,i} = \hat{\mathbf{X}}_{k|k-1}^n + \sqrt{\mathbf{P}_{k|k-1}^n}\xi^i, \quad i = 1, \cdots, 2m \tag{26}$$

Propagate $\bar{\chi}_{k|k-1}^{n,i}$ through nonlinear measurement function to generate sigma points of predicted measurement as

$$\zeta_{k|k-1}^{n,i} = h(\bar{\chi}_{k|k-1}^{n,i}), \quad i = 1, \cdots, 2m \tag{27}$$

The predicted measurement $\hat{\mathbf{Z}}_{k|k-1}^n$ and its error covariance matrix $\mathbf{P}_{z,k|k-1}^n$ are defined by

$$\hat{\mathbf{Z}}_{k|k-1}^n = \frac{1}{2m}\sum_{i=1}^{2m}\zeta_{k|k-1}^{n,i} \tag{28}$$

$$\mathbf{P}_{z,k|k-1}^n = \frac{1}{2m}\sum_{i=1}^{2m}\left(\zeta_{k|k-1}^{n,i} - \hat{\mathbf{Z}}_{k|k-1}^n\right)\left(\zeta_{k|k-1}^{n,i} - \hat{\mathbf{Z}}_{k|k-1}^n\right)^T + \mathbf{R}_k^n \tag{29}$$

where $\mathbf{R}_k^n$ is error covariance matrix of measurement noise at time instant k. The cross-covariance of predicted state and measurement is defined by

$$\mathbf{P}_{xz,k|k-1}^n = \frac{1}{2m}\sum_{i=1}^{2m}\left(\chi_{k|k-1}^{n,i} - \hat{\mathbf{X}}_{k|k-1}^n\right)\left(\zeta_{k|k-1}^{n,i} - \hat{\mathbf{Z}}_{k|k-1}^n\right)^T \tag{30}$$

Error compensation matrix $\mathbf{E}_k^n$ is constructed to offset the impact of neglecting one-step prediction errors of measurement on precision of state estimation [29], and the expression of $\mathbf{E}_k^n$ is given by

$$\mathbf{E}_k^n = \mathbf{P}_{z,k|k-1}^n - \mathbf{H}_k^n\left(\mathbf{Y}_{k|k-1}^n\right)^{-1}\left(\mathbf{H}_k^n\right)^T \tag{31}$$

where $\mathbf{H}_k^n = \left(\mathbf{P}_{xz,k|k-1}^n\right)^T\mathbf{Y}_{k|k-1}^n$ is the pseudo-measurement matrix. The residual error item in (9) is further written by

$$\bar{\zeta}_k^n = \left(\mathbf{E}_k^n\right)^{-1/2}\left(\hat{\mathbf{Z}}_{k|k-1}^n - \mathbf{Z}_k^n\right) \tag{32}$$

where $\bar{\zeta}_k^n = \left[\bar{\zeta}_{k,1}^n, \cdots, \bar{\zeta}_{k,l}^n\right]^T$. Eq. (13) and (14) are updated, respectively, by

$$\psi_i^* = \frac{1}{\sigma^2}\exp\left(-\frac{\left(\bar{\zeta}_{k,i}^n\right)^2}{2\sigma^2}\right), \quad i = 1, \cdots, l \tag{33}$$

$$\psi^* = diag[\psi_1^* \cdots \psi_l^*] \tag{34}$$

The error covariance matrix of measurement noise is defined by

$$\bar{\mathbf{R}}_k^n = \left(\mathbf{E}_k^n\right)^{1/2}\left(\psi^*\right)^{-1}\left(\mathbf{E}_k^n\right)^{T/2} \tag{35}$$

The information contribution equations are defined by

$$\begin{cases} \mathbf{i}_k^n = \left(\mathbf{H}_k^n\right)^T(\bar{\mathbf{R}}_k^n)^{-1}\left[\mathbf{Z}_k^n - \hat{\mathbf{Z}}_{k|k-1}^n + \mathbf{H}_k^n\hat{\mathbf{X}}_{k|k-1}^n\right] \\ \mathbf{I}_k^n = \left(\mathbf{H}_k^n\right)^T(\bar{\mathbf{R}}_k^n)^{-1}\mathbf{H}_k^n \end{cases} \tag{36}$$

Then, the information state vector $\hat{\mathbf{y}}_k^n$ and information matrix $\mathbf{Y}_k^n$ are given by

$$\begin{cases} \hat{\mathbf{y}}_k^n = \hat{\mathbf{y}}_{k|k-1}^n + \mathbf{i}_k^n \\ \mathbf{Y}_k^n = \mathbf{Y}_{k|k-1}^n + \mathbf{I}_k^n \end{cases} \tag{37}$$

Finally, the state estimation at time instant k is defined by

$$\hat{\mathbf{X}}_k^n = \left(\mathbf{Y}_k^n\right)^{-1}\hat{\mathbf{y}}_k^n \tag{38}$$

Remark 1: In practice, there are extremely large outliers caused by the failure of measuring instrument. In this situation, there is no exact solution for $\bar{\mathbf{R}}_k^n$ when the measurement $\mathbf{Z}_k^n \to \infty$. Nevertheless, the value of $\left(\bar{\mathbf{R}}_k^n\right)^{-1}$ is the key to calculate information contribution equations. Hence, denote $\tilde{\mathbf{R}}_k^n = \left(\bar{\mathbf{R}}_k^n\right)^{-1}$, then $\tilde{\mathbf{R}}_k^n = \left(\mathbf{E}_k^n\right)^{-T/2}\psi^*\left(\mathbf{E}_k^n\right)^{-1/2}$. While $\mathbf{Z}_k^n \to \infty$, $\tilde{\mathbf{R}}_k^n \to 0$. Consequently, information contribution equations $\mathbf{i}_k^n \to 0$, $\mathbf{I}_k^n \to 0$ and $\hat{\mathbf{y}}_k^n = \hat{\mathbf{y}}_{k|k-1}^n$, $\mathbf{Y}_k^n = \mathbf{Y}_{k|k-1}^n$.

To cope with extremely large outliers of measurement, (35) and (36) are redefined, respectively, by

$$\tilde{\mathbf{R}}_k^n = \left(\mathbf{E}_k^n\right)^{-T/2}\psi^*\left(\mathbf{E}_k^n\right)^{-1/2} \quad (39)$$

$$\begin{cases}\mathbf{i}_k^n = \left(\mathbf{H}_k^n\right)^T\tilde{\mathbf{R}}_k^n\left[\mathbf{Z}_k^n - \hat{\mathbf{Z}}_{k|k-1}^n + \mathbf{H}_k^n\hat{\mathbf{X}}_{k|k-1}^n\right]\\ \mathbf{I}_k^n = \left(\mathbf{H}_k^n\right)^T\tilde{\mathbf{R}}_k^n\mathbf{H}_k^n\end{cases} \quad (40)$$

IV. WAC BASED DISTRIBUTED ROBUST CIF

This section focuses on the distributed information fusion of distributed WSN. RCIF is employed in the local node and WAC based RCIF is derived for distributed information fusion to improve the stability of distributed WSN and the accuracy of state estimation. Let t be the consensus time. For node n, information pairs $\left(\hat{\mathbf{y}}_k^n, \mathbf{Y}_k^n\right)$ are supposed already known at time k. The consensus weighted factor is defined by $\pi^{n,j}$, where $j \in \mathbb{N}_n$, $\pi^{n,j} \geqslant 0$, $\sum\limits_{j\in N_n}\pi^{n,j} = 1$. The information pair in different nodes has the same behavior as

$$\left(\hat{\mathbf{y}}_k^*, \mathbf{Y}_k^*\right) = \lim_{t\to\infty}\left(\hat{\mathbf{y}}_{k,t}^n, \mathbf{Y}_{k,t}^n\right) \quad (41)$$

The WAC update of information pairs is defined by

$$\begin{cases}\hat{\mathbf{y}}_{k,t+1}^n = \sum\limits_{j\in N_n}\pi^{n,j}\hat{\mathbf{y}}_{k,t}^j\\ \mathbf{Y}_{k,t+1}^n = \sum\limits_{j\in N_n}\pi^{n,j}\mathbf{Y}_{k,t}^j\end{cases} \quad (42)$$

If $\tilde{\mathbf{R}}_k^n = 0$, there are extremely large outliers at node n. Then, node n does not communicate with its adjacent nodes, and consensus weighted value $\pi^{n,j}$ is modified for information fusion in distributed WSN. In other words, node n does not take part in the consensus step at time k. If $\tilde{\mathbf{R}}_k^n \neq 0$, let $\bar{T}$ be the number of WAC iterations, where $t \in [0, \bar{T}-1]$. For node n, the initial information pair is given by $\begin{cases}\hat{\mathbf{y}}_{k,0}^n = \hat{\mathbf{y}}_k^n\\ \mathbf{Y}_{k,0}^n = \mathbf{Y}_k^n\end{cases}$. Then, the updated information pair is $\left(\hat{\mathbf{y}}_{k,\bar{T}}^n, \mathbf{Y}_{k,\bar{T}}^n\right)$ after $\bar{T}$ consensus iteration. Finally, the updated state estimation $\hat{\mathbf{X}}_k^n$ at time instant k after the information fusion is given by

$$\hat{\mathbf{X}}_k^n = \left(\mathbf{Y}_{k,\bar{T}}^n\right)^{-1}\hat{\mathbf{y}}_{k,\bar{T}}^n \quad (43)$$

The general form of DRCIF is summarized in Table 1.

V. STOCHASTIC BOUNDEDNESS OF ESTIMATION ERRORS

Whether the estimation error is bounded or not under mean-square error is a criterion for the performance of DRCIF. Pseudo-process matrix $\Phi_{k|k-1}^n$ and pseudo-measurement matrix $\mathbf{H}_k^n$ are constructed to simplify the derivation process of error boundedness [11], [14], [27]. Diagonal matrix λ_k^n and μ_k^n are introduced to compensate the linear error of process function and measurement function, respectively. μ_k^n and $\mathbf{H}_k^n$ have been given in (6) and (7). λ_k^n and $\Phi_{k|k-1}^n$ are defined by

$$\begin{cases}\Phi_{k|k-1}^n = \left(\mathbf{P}_{x_{k-1},x_{k|k-1}}^n\right)^T\mathbf{Y}_{k-1}^n\\ \lambda_k^n = diag\left[\lambda_{k,1}^n \quad \lambda_{k,2}^n \quad \cdots \quad \lambda_{k,m}^n\right]\end{cases} \quad (44)$$

TABLE 1. Distributed robust cubature information filtering.

Given $\hat{\mathbf{X}}_{k-1}^n$ and $\mathbf{Y}_{k-1}^n$ for node n
Step 1 : Time Update Conduct (20)-(25) for time update to acquire the predicted information matrix $\mathbf{Y}_{k\|k-1}^n$ and the predicted information state vector $\hat{\mathbf{y}}_{k\|k-1}^n$. **Step 2** : Measurement Update Complete (26)-(28) to acquire the predicted measurement $\hat{\mathbf{Z}}_{k\|k-1}^n$ and construct error compensation matrix $\mathbf{E}_k^n$ through (29)-(31). Then calculate information contribution equations $\mathbf{i}_k^n$ and $\mathbf{I}_k^n$ through (32)-(34), (39) and (40). The information state vector $\hat{\mathbf{y}}_k^n$ and information matrix $\mathbf{Y}_k^n$ are acquired by $\begin{cases}\hat{\mathbf{y}}_k^n = \hat{\mathbf{y}}_{k\|k-1}^n + \mathbf{i}_k^n\\ \mathbf{Y}_k^n = \mathbf{Y}_{k\|k-1}^n + \mathbf{I}_k^n\end{cases}$ **Step 3** : Distributed Information Fusion If $\tilde{\mathbf{R}}_k^n = 0$ in local node n $\hat{\mathbf{X}}_k^n = \left(\mathbf{Y}_{k\|k-1}^n\right)^{-1}\hat{\mathbf{y}}_{k\|k-1}^n$ $\mathbf{Y}_k^n = \mathbf{Y}_{k\|k-1}^n$ Adjust the structure of distributed WSN and redefine consensus weighted value $\pi^{n,j}$ for other nodes. else for $t = 0, 1, \cdots \bar{T}-1$, execute the following consensus steps. a) Broadcast the information pair $\left(\hat{\mathbf{y}}_{k,t}^n, \mathbf{Y}_{k,t}^n\right)$ to its neighbors $j \in N_n$. b) Receive the information pair $\left(\hat{\mathbf{y}}_{k,t}^j, \mathbf{Y}_{k,t}^j\right)$ from its neighbors $j \in N_n$. c) Complete the data fusion of information pairs via $\begin{cases}\hat{\mathbf{y}}_{k,t+1}^n = \sum\limits_{j\in N_n}\pi^{n,j}\hat{\mathbf{y}}_{k,t}^j\\ \mathbf{Y}_{k,t+1}^n = \sum\limits_{j\in N_n}\pi^{n,j}\mathbf{Y}_{k,t}^j\end{cases}$ The updated state estimation and information matrix at time k are $\hat{\mathbf{X}}_k^n = \left(\mathbf{Y}_{k,\bar{T}}^n\right)^{-1}\hat{\mathbf{y}}_{k,\bar{T}}^n$ $\mathbf{Y}_k^n = \mathbf{Y}_{k,\bar{T}}^n$ Return the values of $\hat{\mathbf{X}}_k^n$ and $\mathbf{Y}_k^n$ to step 1 for state estimation at time $k+1$.

where $\mathbf{P}_{x_{k-1},x_{k|k-1}}^n = \frac{1}{2m}\sum\limits_{i=1}^{2m}(\chi_{k-1}^{n,i} - \hat{\mathbf{X}}_{k-1}^n)(\chi_{k|k-1}^{n,i} - \hat{\mathbf{X}}_{k|k-1}^n)^T$. m is the dimension of state vector. The system model is further written by

$$\begin{cases}\mathbf{X}_k = \lambda_{k-1}^n\Phi_{k|k-1}^n\mathbf{X}_{k-1} + \omega_{k-1}\\ \mathbf{Z}_k^n = \mu_k^n\mathbf{H}_k^n\mathbf{X}_k + \upsilon_k^n,\end{cases} \quad n = 1, 2, \cdots, N \quad (45)$$

Based on the linear system model, the predicted information matrix and information contribution equations are given by

$$\begin{cases}\mathbf{Y}_{k|k-1}^n = \Big(\lambda_{k-1}^n\Phi_{k|k-1}^n\left(\mathbf{Y}_{k-1}^n\right)^{-1}\left(\lambda_{k-1}^n\Phi_{k|k-1}^n\right)^T\\ \quad +\mathbf{Q}_{k-1}^n\Big)^{-1}\\ \mathbf{i}_k^n = \left(\mu_k^n\mathbf{H}_k^n\right)^T\tilde{\mathbf{R}}_k^n\mathbf{Z}_k^n\\ \mathbf{I}_k^n = \left(\mu_k^n\mathbf{H}_k^n\right)^T\tilde{\mathbf{R}}_k^n\beta_k^n\mathbf{H}_k^n\end{cases} \quad (46)$$

where $\mathbf{Y}_{k|k-1}^n = \left(\mathbf{P}_{k|k-1}^n\right)^{-1}$, $\tilde{\mathbf{R}}_k^n$ is defined in (39). Substitute (6) into (30) to obtain the equivalent form of compensation matrix as

$$\begin{aligned}\mathbf{E}_k^n &= \mathbf{P}_{z,k|k-1}^n - \mathbf{H}_k^n\mathbf{P}_{k|k-1}^n\left(\mathbf{H}_k^n\right)^T\\ &= \mathbf{P}_{z,k|k-1}^n - \left(\mathbf{P}_{xz,k|k-1}^n\right)^T\mathbf{Y}_{k|k-1}^n\mathbf{P}_{k|k-1}^n\end{aligned}$$

$$\begin{aligned}&\times \mathbf{Y}^n_{k|k-1}\mathbf{P}^n_{xz,k|k-1}\\ &= \mathbf{P}^n_{z,k|k-1} - \left(\mathbf{P}^n_{xz,k|k-1}\right)^T \mathbf{Y}^n_{k|k-1}\mathbf{P}^n_{xz,k|k-1}\\ &= \mu^n_k\mathbf{H}^n_k\mathbf{P}^n_{k|k-1}(\mu^n_k\mathbf{H}^n_k)^T + \mathbf{R}^n_k\\ &\quad - \left(\mathbf{P}^n_{k|k-1}(\mu^n_k\mathbf{H}^n_k)^T\right)^T \mathbf{Y}^n_{k|k-1}\mathbf{P}^n_{k|k-1}(\mu^n_k\mathbf{H}^n_k)^T\\ &= \mu^n_k\mathbf{H}^n_k\mathbf{P}^n_{k|k-1}(\mu^n_k\mathbf{H}^n_k)^T + \mathbf{R}^n_k - \mu^n_k\mathbf{H}^n_k\mathbf{P}^n_{k|k-1}(\mu^n_k\mathbf{H}^n_k)^T\\ &= \mathbf{R}^n_k \end{aligned} \tag{47}$$

where $\mathbf{R}^n_k$ is the initial covariance matrix of measurement noise.

Lemma 1: A necessary and sufficient condition for positive matrix $\mathbf{A}$ is that there is a nonsingular matrix $\mathbf{U}$ which meets the equality $\mathbf{A} = \mathbf{U}^T\mathbf{U}$.

Lemma 2: If matrix $\mathbf{A}$ is positive, there is a real number $\underline{A} > 0$ which meets the inequality $\mathbf{A} \geqslant \underline{A}\mathbf{I}_l$, where l is the dimension of $\mathbf{A}$.

Lemma 3 ([27]): Suppose that ε_k is a stochastic process. If a stochastic function $V(\varepsilon_k)$, real number $v_{\min}, v_{\max} > 0$, $\gamma > 0$ and $0 < \eta < 1$ exist such that

$$\begin{cases} v_{\min}\|\varepsilon_k\|^2 \leqslant V(\varepsilon_k) \leqslant v_{\max}\|\varepsilon_k\|^2\\ E\{V(\varepsilon_k)|\varepsilon_{k-1}\} - V(\varepsilon_{k-1}) \leqslant \gamma - \eta V(\varepsilon_{k-1}) \end{cases} \tag{48}$$

are fulfilled for all time instant k. Then the stochastic process is exponentially bounded in mean square, that is

$$E\left\{\|\varepsilon_k\|^2\right\} \leqslant \frac{v_{\max}}{v_{\min}} E\left\{\|\varepsilon_0\|^2\right\}(1-\eta)^k + \frac{\mu}{v_{\min}}\sum_{i=1}^{k-1}(1-\eta)^i \tag{49}$$

Lemma 4 ([30]): If integration $N \geqslant 2$, positive matrices $\mathbf{M}_1, \mathbf{M}_2, \cdots \mathbf{M}_N$ and vectors $\mathbf{M}_1, \mathbf{M}_2, \cdots \mathbf{M}_N$ meet the following inequality

$$\left(\sum_{i=1}^N \mathbf{M}_i\mathbf{v}_i\right)^T \left(\sum_{i=1}^N \mathbf{M}_i\right)^{-1} \left(\sum_{i=1}^N \mathbf{M}_i\mathbf{v}_i\right) \leqslant \sum_{i=1}^N (\mathbf{v}_i)^T\mathbf{M}_i\mathbf{v}_i \tag{50}$$

Assumption 1: Weight matrix $\pi^{\bar{T}} = \left(\pi^{n,j}_{\bar{T}}\right)_{m\times m}$ of WAC is supposed row stochastic and primitive at each time instant k, which can make sure the effectiveness of weight matrix.

Assumption 2: For time instant $k \geqslant 0$, there are real numbers $\bar{\lambda}, \bar{\phi}, \bar{\mu}, \bar{H} \neq 0$ and $\underline{\lambda}, \underline{\phi}, \underline{\mu}, \underline{H} \neq 0$. Then the following inequalities are defined by

$$\begin{cases} \underline{\lambda}^2\mathbf{I}_m \leqslant \lambda^n_k(\lambda^n_k)^T \leqslant \bar{\lambda}^2\mathbf{I}_m\\ \underline{\phi}^2\mathbf{I}_m \leqslant \Phi^n_{k|k-1}\left(\Phi^n_{k|k-1}\right)^T \leqslant \bar{\phi}^2\mathbf{I}_m\\ \underline{\mu}^2\mathbf{I}_l \leqslant \mu^n_k(\mu^n_k)^T \leqslant \bar{\mu}^2\mathbf{I}_l\\ \underline{H}^2\mathbf{I}_l \leqslant \mathbf{H}^n_k(\mathbf{H}^n_k)^T \leqslant \bar{H}^2\mathbf{I}_l \end{cases} \tag{51}$$

Assumption 3: For real number $\bar{q} \geqslant \underline{q} > 0$, $\bar{r} \geqslant \underline{r} > 0$, $p_{\max} \geqslant p_{\min} > 0$, $\bar{p} \geqslant \underline{p} > 0$, the following inequalities are fulfilled

$$\begin{cases} \underline{q}\mathbf{I}_m \leqslant \mathbf{Q}^n_k \leqslant \bar{q}\mathbf{I}_m\\ \underline{r}\mathbf{I}_l \leqslant \mathbf{R}^n_k \leqslant \bar{r}\mathbf{I}_l\\ p_{\min} \leqslant p^n \leqslant p_{\max}\\ \underline{p}\mathbf{I}_m \leqslant \mathbf{P}^n_{k|k-1} \leqslant \bar{p}\mathbf{I}_m \end{cases} \tag{52}$$

Remark 2: The following proof is provided to verify the boundedness of estimation errors for the proposed DRCIF according to Lemma 3. Although the initial covariance of measurement noise is supposed to be bounded, the boundedness of $\tilde{\mathbf{R}}^n_k$ is unknown. Hence $\tilde{\mathbf{R}}^n_k$ must be firstly proved to be bounded, which is a key element to the course of the proof.

Proof: First of all, the local node that generating extremely large outliers of measurement does not take part in the process of information fusion, and the positive definition of $\bar{\mathbf{R}}^n_k$ for other nodes must be proved. Kernel function $\psi^* = (\psi^*)^{1/2}(\psi^*)^{T/2}$ for MCC is a positive matrix. Substitute (47) into (35) and $\bar{\mathbf{R}}^n_k$ is rewritten by

$$\begin{aligned}\bar{\mathbf{R}}^n_k &= (\mathbf{R}^n_k)^{1/2}(\psi^*)^{-1}(\mathbf{R}^n_k)^{T/2}\\ &= (\mathbf{R}^n_k)^{1/2}\left((\psi^*)^{1/2}(\psi^*)^{T/2}\right)^{-1}(\mathbf{R}^n_k)^{T/2}\\ &= \left((\mathbf{R}^n_k)^{1/2}(\psi^*)^{-T/2}\right)\left((\psi^*)^{-1/2}(\mathbf{R}^n_k)^{T/2}\right) \end{aligned} \tag{53}$$

Denote $\mathbf{U} = (\psi^*)^{-1/2}(\mathbf{R}^n_k)^{T/2}$. Because of the positive definition of ψ^* and $\mathbf{R}^n_k$, $\mathbf{U}$ is a positive matrix. Then, matrix $\bar{\mathbf{R}}^n_k = \mathbf{U}^T\mathbf{U}$ is positive according to Lemma 1. Based on Lemma 2, there is a real number $\underline{r}' > 0$ which meets the inequality $\bar{\mathbf{R}}^n_k \geqslant \underline{r}'\mathbf{I}_l$, and then $\tilde{\mathbf{R}}^n_k = \left(\bar{\mathbf{R}}^n_k\right)^{-1} \leqslant \left(\underline{r}'\right)^{-1}\mathbf{I}_l$.

Secondly, the stochastic boundedness of estimation errors is proved according to Lemma 3. For node n, the predicted error $\tilde{\mathbf{x}}^n_{k|k-1}$ and estimation error $\tilde{\mathbf{x}}^n_k$ are defined by $\tilde{\mathbf{x}}^n_{k|k-1} = \mathbf{x}_k - \hat{\mathbf{x}}^n_{k|k-1}$ and $\tilde{\mathbf{x}}^n_k = \mathbf{x}_k - \hat{\mathbf{x}}^n_k$, respectively. Let $\tilde{\mathbf{x}}_{k|k-1} = col\left(\tilde{\mathbf{x}}^n_{k|k-1}, n \in N\right)$ be the column set of $\tilde{\mathbf{x}}^n_{k|k-1}$ and $\tilde{\mathbf{x}}_k = col\left(\tilde{\mathbf{x}}^n_k, n \in N\right)$ be the column set of $\tilde{\mathbf{x}}^n_k$. Let $\mathbf{p} = \left(p^1, \cdots, p^n, \cdots, p^N\right)^T$ be the Perron-Frobenius left eigenvector of weight matrix $\pi^{\bar{T}}$. p^n is positive according to assumption 3. Vector $\mathbf{p}$ meets the equality $\mathbf{p}^T\pi^{\bar{T}} = \mathbf{p}^T$, which equals to $\sum_{j\in N} p^n\pi^{n,j}_{\bar{T}} = p^n$. Construct a stochastic process through $\tilde{\mathbf{x}}^n_{k|k-1}$

$$V\left(\tilde{\mathbf{x}}^n_{k|k-1}\right) = \sum_{n\in N} p^n\left(\tilde{\mathbf{x}}^n_{k|k-1}\right)^T \mathbf{Y}^n_{k|k-1}\tilde{\mathbf{x}}^n_{k|k-1} \tag{54}$$

where $\mathbf{Y}^n_{k|k-1}$ is given in (46). According to assumption 2 and assumption 3, the stochastic bound of $\mathbf{Y}^n_{k|k-1}$ is defined by

$$\left(\bar{\lambda}^2\bar{\phi}^2\bar{p} + \bar{q}\right)^{-1}\mathbf{I}_m \leqslant \mathbf{Y}^n_{k|k-1} \leqslant \left(\underline{\lambda}^2\underline{\phi}^2\underline{p} + \underline{q}\right)^{-1}\mathbf{I}_m \tag{55}$$

Denote $\left(\bar{\lambda}^2\bar{\phi}^2\bar{p}+\bar{q}\right)^{-1} = \underline{Y}$ and $\left(\underline{\lambda}^2\underline{\phi}^2\underline{p}+\underline{q}\right)^{-1} = \bar{Y}$, (55) is further written by

$$\underline{Y}\mathbf{I}_m \leqslant \mathbf{Y}^n_{k|k-1} \leqslant \bar{Y}\mathbf{I}_m \tag{56}$$

Then, the stochastic bound for $V\left(\tilde{\mathbf{x}}^n_{k|k-1}\right)$ is

$$p_{\min}\underline{Y}\left\|\tilde{\mathbf{x}}^n_{k|k-1}\right\|^2 \leqslant V\left(\tilde{\mathbf{x}}^n_{k|k-1}\right) \leqslant p_{\max}\bar{Y}\left\|\tilde{\mathbf{x}}^n_{k|k-1}\right\|^2 \tag{57}$$

which meets the first term of (48) on application of Lemma 3. To prove the second term of (48), the predicted error $\tilde{\mathbf{x}}^n_{k|k-1}$ is transformed into

$$\begin{aligned}\tilde{\mathbf{x}}^n_{k|k-1} &= \mathbf{x}_k - \hat{\mathbf{x}}^n_{k|k-1}\\ &= \lambda^n_{k-1}\Phi^n_{k|k-1}\left(\mathbf{x}_{k-1}-\hat{\mathbf{x}}^n_{k-1}\right)+\omega^n_{k-1}\\ &= \lambda^n_{k-1}\Phi^n_{k|k-1}\sum_{j\in N}\pi^{n,j}_{\bar{T}}\Big[\left(\mathbf{Y}^n_{k-1}\right)^{-1}\mathbf{Y}^j_{k-1|k-2}\tilde{\mathbf{x}}^j_{k-1|k-2}\\ &\quad -\left(\mathbf{Y}^n_{k-1}\right)^{-1}\left(\mu^j_{k-1}\mathbf{H}^j_{k-1}\right)^T\tilde{\mathbf{R}}^n_k\upsilon^j_{k-1}\Big]+\omega^n_{k-1}\\ &= \sum_{j\in N}\Upsilon^{n,j}_{k-1}\tilde{\mathbf{x}}^n_{k-1|k-2}+\sum_{j\in N}\Xi^{n,j}_{k-1}\upsilon^n_{k-1}+\omega^n_{k-1}\end{aligned} \tag{58}$$

where

$$\begin{cases}\Upsilon^{n,j}_{k-1} = \lambda^n_{k-1}\pi^{n,j}_{\bar{T}}\Phi^n_{k|k-1}\left(\mathbf{Y}^n_{k-1}\right)^{-1}\mathbf{Y}^j_{k-1|k-2}\\ \Xi^{n,j}_{k-1} = -\pi^{n,j}_{\bar{T}}\lambda^n_{k-1}\Phi^n_{k|k-1}\left(\mathbf{Y}^n_{k-1}\right)^{-1}\left(\mu^j_{k-1}\mathbf{H}^j_{k-1}\right)^T\tilde{\mathbf{R}}^n_k\end{cases} \tag{59}$$

Insert (59) into (54) and take conditional expectation of (54), one can obtain

$$E\left\{V\left(\tilde{\mathbf{x}}_{k|k-1}\right)|\tilde{\mathbf{x}}_{k-1|k-2}\right\} = \Lambda^x_{k-1}+\Lambda^\xi_{k-1}+\Lambda^\omega_{k-1} \tag{60}$$

where

$$\begin{cases}\Lambda^x_{k-1} = E\Bigg\{\sum_{n\in N}p^n\left(\sum_{j\in N}\Upsilon^{n,j}_{k-1}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)^T\mathbf{Y}^n_{k|k-1}\\ \quad\times\left(\sum_{j\in N}\Upsilon^{n,j}_{k-1}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\Bigg\}\\ \Lambda^\xi_{k-1} = E\Bigg\{\sum_{n\in N}p^n\left(\sum_{j\in N}\Xi^{n,j}_{k-1}\upsilon^j_{k-1}\right)^T\mathbf{Y}^n_{k|k-1}\\ \quad\times\left(\sum_{j\in N}\Xi^{n,j}_{k-1}\upsilon^j_{k-1}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\Bigg\}\\ \Lambda^\omega_{k-1} = E\left\{\sum_{n\in N}p^n\left(\omega^n_{k-1}\right)^T\mathbf{Y}^n_{k|k-1}\left(\omega^n_{k-1}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\right\}\end{cases} \tag{61}$$

The upper bound of Λ^x_{k-1} is considered. Denote an instant value $0 < \bar{\kappa} < 1$ and $\mathbf{Y}^n_{k|k-1}$ meets the following inequality [14]

$$\begin{aligned}\mathbf{Y}^n_{k|k-1} &= \left(\lambda^n_{k-1}\Phi^n_{k|k-1}\left(\mathbf{Y}^n_{k-1}\right)^{-1}\left(\lambda^n_{k-1}\Phi^n_{k|k-1}\right)^T+\mathbf{Q}^n_{k-1}\right)^{-1}\\ &\leqslant \bar{\kappa}\left(\lambda^n_{k-1}\Phi^n_{k|k-1}\right)^{-T}\mathbf{Y}^n_{k-1}\left(\lambda^n_{k-1}\Phi^n_{k|k-1}\right)^{-1}\end{aligned} \tag{62}$$

Substitute (62) into the first term of (61), one can obtain

$$\begin{aligned}\Lambda^x_{k-1} &\leqslant \bar{\kappa}E\Bigg\{\sum_{n\in N}p^n\\ &\quad\times\left(\sum_{j\in N}\Upsilon^{n,j}_{k-1}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)^T\left(\lambda^n_{k-1}\Phi^n_{k|k-1}\right)^{-T}\mathbf{Y}^n_{k-1}\\ &\quad\times\left(\lambda^n_{k-1}\Phi^n_{k|k-1}\right)\left(\sum_{j\in N}\Upsilon^{n,j}_{k-1}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\Bigg\}\\ &= \bar{\kappa}E\Bigg\{\sum_{n\in N}p^n\left(\sum_{j\in N}\pi^{n,j}_{\bar{T}}\mathbf{Y}^j_{k-1|k-2}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)^T\\ &\quad\times\left(\mathbf{Y}^n_{k-1}\right)^{-1}\left(\sum_{j\in N}\pi^{n,j}_{\bar{T}}\mathbf{Y}^j_{k-1|k-2}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\Bigg\}\end{aligned} \tag{63}$$

Since $\mathbf{Y}^n_{k-1} \geqslant \sum_{j\in N}\pi^{n,j}_{\bar{T}}\mathbf{Y}^j_{k-1|k-2}$, (63) is further written by

$$\begin{aligned}\Lambda^x_{k-1} &\leqslant \bar{\kappa}E\Bigg\{\sum_{n\in N}p^n\\ &\quad\times\left(\sum_{j\in N}\pi^{n,j}_{\bar{T}}\mathbf{Y}^j_{k-1|k-2}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)^T\left(\sum_{j\in N}\pi^{n,j}_{\bar{T}}\mathbf{Y}^j_{k-1|k-2}\right)^{-1}\\ &\quad\times\left(\sum_{j\in N}\pi^{n,j}_{\bar{T}}\mathbf{Y}^j_{k-1|k-2}\tilde{\mathbf{x}}^j_{k-1|k-2}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\Bigg\}\end{aligned} \tag{64}$$

According to Lemma 4, the upper bound of Λ^x_{k-1} is defined by

$$\begin{aligned}\Lambda^x_{k-1} &\leqslant \bar{\kappa}E\Bigg\{\sum_{n\in N}p^n\sum_{j\in N}\pi^{n,j}_{\bar{T}}\left(\tilde{\mathbf{x}}^j_{k-1|k-2}\right)^T\mathbf{Y}^j_{k-1|k-2}\\ &\quad\times\left(\tilde{\mathbf{x}}^j_{k-1|k-2}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\Bigg\}\\ &= \bar{\kappa}E\Bigg\{\sum_{j\in N}p^j\left(\tilde{\mathbf{x}}^j_{k-1|k-2}\right)^T\mathbf{Y}^j_{k-1|k-2}\\ &\quad\times\left(\tilde{\mathbf{x}}^j_{k-1|k-2}\right)\Big|\tilde{\mathbf{x}}_{k-1|k-2}\Bigg\}\\ &= \bar{\kappa}V\left(\tilde{\mathbf{x}}_{k-1|k-2}\right)\end{aligned} \tag{65}$$

Then, the upper bound of $\Lambda_{k-1}^{\xi}+\Lambda_{k-1}^{\omega}$ is proved.

$$
\begin{aligned}
&\Lambda_{k-1}^{\xi}+\Lambda_{k-1}^{\omega}\\
&=E\left\{\sum_{n\in N}p^{n}\left[\left(\sum_{j\in N}\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)^{T}\mathbf{Y}_{k|k-1}^{n}\left(\sum_{j\in N}\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)\right.\right.\\
&\qquad\left.\left.+\left(\omega_{k-1}^{n}\right)^{T}\mathbf{Y}_{k|k-1}^{n}\left(\omega_{k-1}^{n}\right)\right]\Big|\tilde{\mathbf{x}}_{k-1|k-2}\right\}\\
&\leqslant\bar{Y}E\left\{\sum_{n\in N}p^{n}\left[\left(\sum_{j\in N}\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)^{T}\left(\sum_{j\in N}\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)\right.\right.\\
&\qquad\left.\left.+\left(\omega_{k-1}^{n}\right)^{T}\left(\omega_{k-1}^{n}\right)\right]\Big|\tilde{\mathbf{x}}_{k-1|k-2}\right\}\\
&=\bar{Y}E\left\{\sum_{n\in N}p^{n}\left[\sum_{j\in N}\left(\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)^{T}\left(\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)\right.\right.\\
&\qquad\left.\left.+\left(\omega_{k-1}^{n}\right)^{T}\left(\omega_{k-1}^{n}\right)\right]\Big|\tilde{\mathbf{x}}_{k-1|k-2}\right\}\\
&=\bar{Y}E\left\{\sum_{n\in N}p^{n}\left[tr\left(\sum_{j\in N}\left(\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)^{T}\left(\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)\right)\right.\right.\\
&\qquad\left.\left.+tr\left(\left(\omega_{k-1}^{n}\right)^{T}\left(\omega_{k-1}^{n}\right)\right)\right]\Big|\tilde{\mathbf{x}}_{k-1|k-2}\right\}\\
&=\bar{Y}E\left\{\sum_{n\in N}p^{n}\left[\sum_{j\in N}tr\left(\left(\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)^{T}\left(\Xi_{k-1}^{n,j}\upsilon_{k-1}^{j}\right)\right)\right.\right.\\
&\qquad\left.\left.+tr\left(\left(\omega_{k-1}^{n}\right)^{T}\left(\omega_{k-1}^{n}\right)\right)\right]\Big|\tilde{\mathbf{x}}_{k-1|k-2}\right\}\\
&=\bar{Y}\left\{\sum_{n\in N}p^{n}\left[\sum_{j\in N}tr\left(\left(\Xi_{k-1}^{n,j}\right)^{T}\left(\Xi_{k-1}^{n,j}\mathbf{R}_{k-1}^{n}\right)\right)\right.\right.\\
&\qquad\left.\left.+tr\left(\mathbf{Q}_{k-1}^{n}\right)\right]\Big|\tilde{\mathbf{x}}_{k-1|k-2}\right\}
\end{aligned}\tag{66}
$$

According to assumption 2 and assumption 3 and substitute $\Xi_{k-1}^{n,j}$ of (59) into (66), the upper bound of $\Lambda_{k-1}^{\xi}+\Lambda_{k-1}^{\omega}$ is defined by

$$
\begin{aligned}
&\Lambda_{k-1}^{\xi}+\Lambda_{k-1}^{\omega}\\
&\qquad\leqslant\bar{Y}\sum_{n\in N}p^{n}\left[\sum_{j\in N}\left(\pi_{\bar{T}}^{n,j}\right)^{2}\frac{\bar{\lambda}^{2}\bar{\phi}^{2}\bar{p}^{2}\bar{\mu}^{2}\bar{H}^{2}\bar{r}l}{\left(\underline{r}'\right)^{2}}+\bar{q}m\right]\triangleq\gamma
\end{aligned}\tag{67}
$$

Combine (65) and (67), one can obtain

$$
\begin{aligned}
E\left\{V\left(\tilde{\mathbf{x}}_{k|k-1}\right)|\tilde{\mathbf{x}}_{k-1|k-2}\right\}-V\left(\tilde{\mathbf{x}}_{k-1|k-2}\right)&\\
\leqslant\gamma-\eta V\left(\tilde{\mathbf{x}}_{k-1|k-2}\right)&
\end{aligned}\tag{68}
$$

where $\eta=1-\bar{\kappa}$, $0<\eta<1$. Eq. (68) meets the second condition of (48) on the application of Lemma 3. The predicted $\tilde{\mathbf{x}}_{k|k-1}$ is proved to be exponentially bounded in mean square.

Finally, the exponential bound of estimation error $\tilde{\mathbf{x}}_{k}^{n}$ is proved. $\tilde{\mathbf{x}}_{k}^{n}$ and $\tilde{\mathbf{x}}_{k+1|k}^{n}$ meet the following equality

$$
\tilde{\mathbf{x}}_{k+1|k}^{n}=\lambda_{k}^{n}\Phi_{k+1|k}^{n}\left(\mathbf{x}_{k}-\hat{\mathbf{x}}_{k}^{n}\right)+\omega_{k}^{n}\tag{69}
$$

Calculate the expectation of (69) and the following inequality is fulfilled.

$$
E\left(\left\|\tilde{\mathbf{x}}_{k}^{n}\right\|^{2}\right)\leqslant\left(\underline{\lambda}\underline{\phi}\right)^{-2}\left(E\left(\left\|\tilde{\mathbf{x}}_{k+1|k}^{n}\right\|^{2}\right)-E\left(\left\|\omega_{k}^{n}\right\|^{2}\right)\right)\tag{70}
$$

Employing the same technique as before, ω_{k}^{n} still can be proved exponentially bounded in mean square. Therefore, $\tilde{\mathbf{x}}_{k}^{n}$ is exponentially bounded in mean square. The proof is completed.

Remark 3: To simplify the proof process, we select linearization approximation [11], [14], [31] of nonlinear system models to analyse the boundedness of estimation errors. Although the compensation vectors λ_{k}^{n} and μ_{k}^{n} is related to nonlinear system model in (1), it is unnecessary to obtain the exact magnitude of λ_{k}^{n} and μ_{k}^{n}.

VI. PERFORMANCE EVALUATION AND DISCUSSION

To evaluate the effectiveness of the proposed algorithm, MCC based unscented Kalman filtering (MCC-UKF) [23], MCC based cubature Kalman filtering (MCC-CKF) [25] and WAC based distributed cubature information filtering (DCIF) [14] are selected for comparison analysis. Generally, the Gaussian kernel is selected as $G_{\sigma}=\exp(-\frac{e^{2}}{2\sigma^{2}})$ in MCC-UKF and MCC-CKF, but in this paper, the kernel function is derived as $\psi=\frac{1}{\sigma^{2}}\exp\left(-\frac{e^{2}}{2\sigma^{2}}\right)$. To verify the impact of kernel bandwidth as coefficients on the accuracy of state estimation, a new DRCIF (DRCIF1) is designed with kernel function $\psi=\exp\left(-\frac{e^{2}}{2\sigma^{2}}\right)$ and the algorithm flow of DRCIF1 is still consistent with DRCIF.

The coordinated turn (CT) model [14] is chosen in the following simulation and a fixed angular speed γ of turning is employed for maneuvering target in X/Y plane. The process equation is given by

$$
\mathbf{X}_{k}=\begin{bmatrix}1&0&\dfrac{\sin(\gamma T)}{\gamma}&\dfrac{-(1-\cos(\gamma T))}{\gamma}\\0&1&\dfrac{1-\cos(\gamma T)}{\gamma}&\dfrac{\sin(\gamma T)}{\gamma}\\0&0&\cos(\gamma T)&-\sin(\gamma T)\\0&0&\sin(\gamma T)&\cos(\gamma T)\end{bmatrix}\mathbf{X}_{k-1}
+\begin{bmatrix}\dfrac{T^{2}}{2}&0\\0&\dfrac{T^{2}}{2}\\T&0\\0&T\end{bmatrix}\omega_{k-1}\tag{71}
$$

where $\mathbf{X}_{k}=\left[p_{k,x},p_{k,y},\dot{p}_{k,x},\dot{p}_{k,y}\right]^{T}$ is the state vector at time instant k. $p_{k,x}$ and $p_{k,y}$ denote the target position. $\dot{p}_{k,x}$ and $\dot{p}_{k,y}$

denote the target speed. $\omega_k \sim N(0, \mathbf{Q}_k)$ is the process noise with zero-mean and covariance $\mathbf{Q}_k$. T is the sampling period of the system.
The measurement equation for distributed WSN is defined by

$$\mathbf{Z}_k^n = \begin{bmatrix} \rho_k^n \\ \theta_k^n \end{bmatrix} = \begin{bmatrix} \sqrt{\left(p_{k,x} - p_x^n\right)^2 + \left(p_{k,y} - p_y^n\right)^2} \\ arc\tan\left(\frac{p_{k,y} - p_y^n}{p_{k,x} - p_x^n}\right) \end{bmatrix} + \upsilon_k^n \tag{72}$$

where $\mathbf{Z}_k^n$ is the measurement vector in sensor node n at time instant k. $\left[p_x^n, p_y^n\right]^T$ denotes the position of node n. ρ_k^n and θ_k^n denote the distance and angle from the target to node n, respectively. $\upsilon_k^n \sim N[0, \mathbf{R}_k]$ is the measurement noise with zero-mean and covariance $\mathbf{R}_k$. Sensor nodes in distributed WSN are employed at the position of (20, 0), (80, 0), (140, 0), (200, 0), (10, 40), (70, 40), (130, 40), (190, 40), (0, 80), (60, 80), (120, 80) and (180, 80). The unit of position is m. The topology map of distributed WSN is displayed in Fig. 1. According to [28], [30], [31], the weighted matrix is defined through Metropolis weight rule.

$$\pi^{n,j} = \begin{cases} 1/\left(1 + \max\left\{d_n, d_j\right\}\right), & if\ (n,j) \in \varepsilon \\ 1 - \sum\limits_{(n,j)\in\varepsilon} \pi^{n,j}, & if\ n = j \\ 0, & \text{otherwise} \end{cases} \tag{73}$$

where d_j is the number of jth node's neighbours.

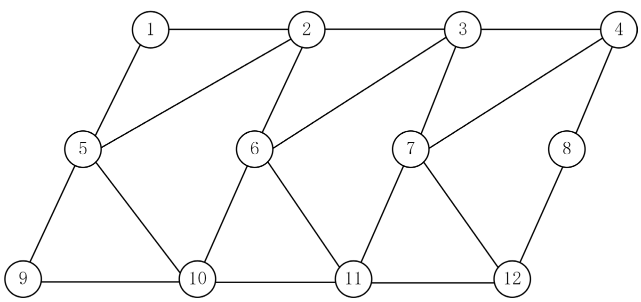


FIGURE 1. **The topology map of distributed WSN.**

In this paper, we mainly investigate the problem of state estimation under measurement outliers or impulse noise, which belongs to the scope of non-Gaussian measurement noise. Hence, three different scenarios of state estimation are designed: a) state estimation under Gaussian noise and impulse noise; b) state estimation under Gaussian mixture noise with heavy-tail property; c) state estimation under Gaussian mixture noise and impulse noise. Sensor nodes are equivalent to each other. MCC-UKF and MCC-CKF are designed for state estimation on a single sensor node, but DCIF, DRCIF1 and DRCIF are derived for state evaluation in distribution WSN. Therefore, the state estimation of node 5 is selected for comparison analysis and 100 independent Mentos Carlo is selected for following simulations. Then root mean square error (RMSE) and average root mean square error (ARMSE) are chosen to verify the effectiveness of DRCIF. RMSE and ARMSE of position are defined, respectively, as

$$\text{RMSE}_{p,k}^n = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\left(p_{k,x}^i - \hat{p}_{k,x}^i\right)^2 + \left(p_{k,y}^i - \hat{p}_{k,y}^i\right)^2\right)} \tag{74}$$

$$\text{ARMSE}_{p,K}^n = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}\left(\left(p_{k,x}^i - \hat{p}_{k,x}^i\right)^2 + \left(p_{k,y}^i - \hat{p}_{k,y}^i\right)^2\right)} \tag{75}$$

where the variable M is the number of Monte Carlo runs. $\left[p_{k,x}, p_{k,y}\right]$ and $\left[\hat{p}_{k,x}, \hat{p}_{k,y}\right]$ stand for the real position and estimated position of the target at time instant k, respectively. K is the simulation period. Initial parameters for filtering algorithms are given as follows. The number of consensus steps is $\bar{T} = 5$ for WAC based algorithms. The simulation period is $K = 100s$. The initial state vector of the target is $\mathbf{X}_0 = \left[100m\ 60m\ 3m/s\ 1m/s\right]^T$ and its corresponding covariance matrix is $\mathbf{P}_0 = diag[25, 16, 0.25, 0.25]$. The angular speed of the target is $\gamma = 2rad/s$. Covariance matrices of process noise and measurement noise are $\mathbf{Q}_k = 0.01 * I_2$ and $\mathbf{R}_k = diag[\,0.25\ 4 * 10^{-4}\,]$, respectively.

A. STATE ESTIMATION UNDER GAUSSIAN NOISE AND IMPULSE NOISE

Impulse noise is injected into measurements at different time instants of 19s, 26s, 27s, 31s, 33s, 57s, 65s, 69s, 87s and 97s with imparity values of $(175.1m\ 1.8rad)^T$, $(60.5m\ 0.2rad)^T$, $(25.8m\ 0.3rad)^T$, $(197.8m\ 1.4rad)^T$, $(75.1m\ 0.1rad)^T$, $(65.1m\ 1.5rad)^T$, $(5m\ 0.5rad)^T$, $(176.3m\ 0.8rad)^T$, $(93.4m\ 0.9rad)^T$ and $(128.7m\ 0.7rad)^T$, respectively. The parameter of Gaussian noise is the same with the initial condition. Different position ARMSEs of DRCIF are presented in Table 2 under various kernel bandwidth σ, values of which are 1, 5, 10 and 100, respectively. Meanwhile, position ARMSEs of MCC-UKF, MCC-CKF, and DRCIF1 are also displayed for comparison analysis. Position RMSEs of related algorithms are presented in figure 2.

From Table 2, although kernel functions are equivalent to each other among MCC-UKF, MCC-CKF and DRCIF1, ARMSE of DRCIF1 is better than the corresponding values of MCC-UKF and MCC-CKF when $\sigma = 1,\ 5,\ 10$ due to the WAC based distributed fusion method for state estimation. However, while $\sigma = 100$, ARMSE of DRCIF1 is larger than the corresponding values of MCC-UKF and MCC-CKF, and we can get that kernel bandwidth with large values is useless to improve the precision of state estimation. The inference is also verified by the trend of ARMSE for MCC-UKF and MCC-CKF. DRCIF acquires better ARMSE than DRCIF1, and one point is obvious that kernel function with coefficients

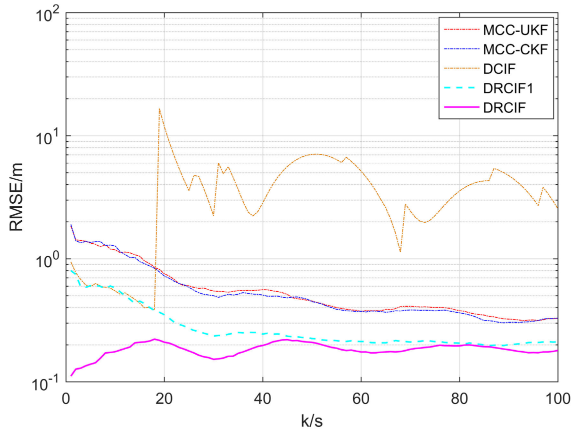


FIGURE 2. Position RMSE under Gaussian noise and impulse noise.

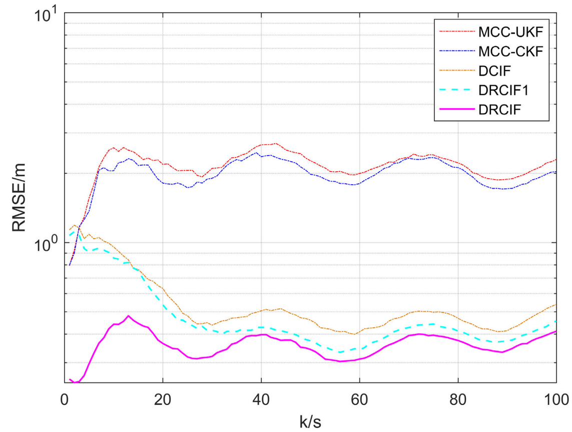


FIGURE 3. Position RMSE under Gaussian mixture noise.

TABLE 2. Position ARMSE for various σ under Gaussian noise and impulse noise.

σ	1	5	10	100
MCC-UKF	0.7256	0.6754	0.6923	0.7129
MCC-CKF	0.7048	0.6746	0.6928	0.7188
DRCIF1	0.3483	0.3146	0.3591	1.3238
DRCIF	0.3140	0.2301	0.2120	0.1418

can obtain better accuracy of state estimation than kernel function without factors. ARMSE of DRCIF decreases with the increase of kernel bandwidth, but the improvement of precision is limited. Hence, $1 < \sigma < 5$ is suggested for the related algorithms and $\sigma = 3$ is selected for the following simulations.

MCC-UKF and MCC-CKF acquire the similar RMSE from Fig. 2. RMSE of DCIF is divergent after $19s$ since DCIF is founded based on MMSE criterion which cannot cope with state estimation under measurement outliers. Benefited from the distributed fusion method based on WAC, DRCIF1 obtains better RMSE than MCC-UKF and MCC-CKF. RMSE of DRCIF is better than DRCIF1 on the account of the added coefficient in kernel function. We can get that DRCIF is more applicable to state estimation under Gaussian noise and impulse noise than the other related algorithms.

B. STATE ESTIMATION UNDER GAUSSIAN MIXTURE NOISE

Gaussian mixture noise with heavy-tail property is considered to evaluate the performance of the proposed robust algorithm, and the definition of measurement noise is given by

$$\upsilon_k^n \sim 0.8N\left(0, diag[\,0.25\ \ 4*10^{-4}\,]\right) + 0.2N\left(0, diag[\,1\ \ 1\,]\right)$$

where 0.8 and 0.2 are probabilities. RMSEs of related algorithms are displayed in Fig. 3.

From Fig. 3, MCC-UKF and MCC-CKF obtain the worst RMSEs since the information fusion of state estimation is not contained in both algorithms. Although the MCC based robust method is not included, RMSE of DCIF is similar to DRCIF1 on account of WAC based information fusion methods. Another point is obvious that the method of information fusion can depress the impact of Gaussian mixture noise on state estimation. Since the best RMSE is obtained, DRCIF is more suitable for state estimation under Gaussian mixture noise than the other related algorithms.

C. STATE ESTIMATION UNDER GAUSSIAN MIXTURE NOISE AND IMPULSE NOISE

In this section, parameters setting of Gaussian mixture noise are consistent with the related items in section 4.2, and the setting of impulse noise stays same with section 4.1. The position RMSEs of related algorithms are displayed in Fig. 4.

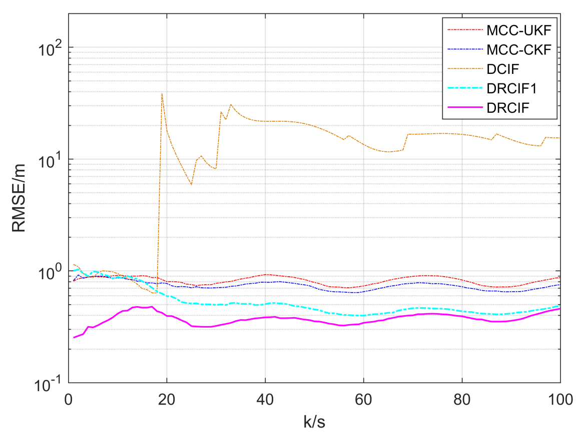


FIGURE 4. Position RMSE under Gaussian mixture noise and impulse noise.

In Fig. 4, RMSE of DCIF is still divergent after $19s$ due to the impact of measurement outliers on state estimation. Benefited from WAC based distributed fusion method, contained robust prosperity and added coefficients in kernel function, DRCIF obtains the best RMSE among the related algorithms. Because of incorporated information fusion process, RMSE of DRCIF1 is better than MCC-UKF and MCC-CKF after $15s$ when the state estimator is stable. Finally, DRCIF is more applicable to state estimation under non-Gaussian measurement noise than the other related algorithms.

VII. CONCLUSION

In this paper, a distributed robust method is designed for nonlinear system to cope with state estimation under measurement outliers or impulse noise in WSN. In the local sensor node, MCC based RCIF is derived to deal with measurement outliers. Meanwhile, covariance matrix of measurement noise is modified to restrain the filtering divergence caused by extremely large outliers. For the distributed WSN, WAC based DRCIF is designed for information fusion to improve the accuracy of state estimation and the stability of distributed networks. Particularly, WAC based information fusion methods belong to the scope of CI, the stability of which can be guaranteed by any times of consensus step (even a single step). Then, the evaluated error of DRCIF is proved to be bounded in mean square. Furthermore, the added coefficient of kernel function in DRCIF can improve the precision of state estimation, which is verified by numerical simulations.

REFERENCES

[1] V. Potdar, A. Sharif, and E. Chang, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, 2002.

[2] Y. Wang, J. Li, and Q. Sun, "Coordinated target tracking by distributed unscented information filter in sensor networks with measurement constraints," *Math. Problems Eng.*, vol. 2013, Dec. 2013, Art. no. 402732.

[3] Q. Gan and C. Harris, "Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 1, pp. 273–279, Jan. 2001.

[4] H. Durrant-Whyte, "A beginners guide to decentralised data fusion," *Saunders Phys. Activities*, vol. 32, no. 1, pp. 333–337, 2006.

[5] D. Li, K. Wong, Y. Hen Hu, and A. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 17–29, Mar. 2002.

[6] Q. Zhang, C. Zhang, M. Liu, and S. Zhang, "Local node selection for target tracking based on underwater wireless sensor networks," *Int. J. Syst. Sci.*, vol. 46, no. 16, pp. 2918–2927, Dec. 2015.

[7] S. Das and J. M. F. Moura, "Consensus+innovations distributed Kalman filter with optimized gains," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 467–481, Jan. 2017.

[8] G. Battistelli and L. J. A. Chisci, "Stability of consensus extended Kalman filtering for distributed state estimation," *Automatica*, vol. 68, no. 3, pp. 169–178, 2016.

[9] X. Hu, B. Xu, and Y. H. Hu, "Target tracking with distance-dependent measurement noise in wireless sensor networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 5200–5203.

[10] A. Keshavarz-Mohammadiyan and H. Khaloozadeh, "Consensus-based distributed unscented target tracking in wireless sensor networks with state-dependent noise," *Signal Process.*, vol. 144, pp. 283–295, Mar. 2018.

[11] W. Li, G. Wei, F. Han, and Y. Liu, "Weighted average consensus-based unscented Kalman filtering," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 558–567, Feb. 2016.

[12] V. P. Bhuvana, M. Schranz, M. Huemer, and B. Rinner, "Distributed object tracking based on cubature Kalman filter," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2013, pp. 423–427.

[13] Q. Chen, C. Yin, J. Zhou, Y. Wang, X. Wang, and C. Chen, "Hybrid consensus-based cubature Kalman filtering for distributed state estimation in sensor networks," *IEEE Sensors J.*, vol. 18, no. 11, pp. 4561–4569, Jun. 2018.

[14] Q. Chen, W. Wang, C. Yin, X. Jin, and J. Zhou, "Distributed cubature information filtering based on weighted average consensus," *Neurocomputing*, vol. 243, pp. 115–124, Jun. 2017.

[15] J. Zhang, S. Gao, Y. Zhong, X. Qi, J. Xia, and J. Yang, "An advanced cubature information filtering for indoor multiple wideband source tracking with a distributed noise statistics estimator," *IEEE Access*, vol. 7, pp. 151851–151866, 2019.

[16] B. Gao, S. Gao, G. Hu, Y. Zhong, and C. Gu, "Maximum likelihood principle and moving horizon estimation based adaptive unscented Kalman filter," *Aerosp. Sci. Technol.*, vol. 73, pp. 184–196, Feb. 2018.

[17] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control.*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.

[18] K. Pakki, B. Chandra, D.-W. Gu, and I. Postlethwaite, "Cubature information filter and its applications," in *Proc. Amer. Control Conf.*, Jun. 2011, pp. 3609–3614.

[19] Y. Chen, Q. Zhao, and R. Liu, "A novel square-root cubature information weighted consensus filter algorithm for distributed camera networks," *Acta Electron. Sinica*, vol. 44, no. 10, pp. 2335–2343, 2016.

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

[21] B. Chen, X. Liu, H. Zhao, and J. C. Principe, "Maximum correntropy Kalman filter," *Automatica*, vol. 76, pp. 70–77, Feb. 2017.

[22] X. Liu, B. Chen, H. Zhao, J. Qin, and J. Cao, "Maximum correntropy Kalman filter with state constraints," *IEEE Access*, vol. 5, pp. 25846–25853, 2017.

[23] X. Liu, H. Qu, J. Zhao, P. Yue, and M. Wang, "Maximum correntropy unscented Kalman filter for spacecraft relative state estimation," *Sensors*, vol. 16, no. 9, p. 1530, Sep. 2016.

[24] G. Wang, Y. Zhang, and X. Wang, "Iterated maximum correntropy unscented Kalman filters for non-Gaussian systems," *Signal Process.*, vol. 163, pp. 87–94, Oct. 2019.

[25] X. Liu, H. Qu, J. Zhao, and P. Yue, "Maximum correntropy square-root cubature Kalman filter with application to SINS/GPS integrated systems," *ISA Trans.*, vol. 80, pp. 195–202, Sep. 2018.

[26] J. Zhang, S. Gao, B. Gao, W. Wei, and J. Yang, "An improved cubature information filtering for target tracking in distributed wireless sensor networks with measurement outliers," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Jun. 2019, pp. 4016–4021.

[27] J. Zarei and E. Shokri, "Convergence analysis of non-linear filtering based on cubature Kalman filter," *IET Sci., Meas. Technol.*, vol. 9, no. 3, pp. 294–305, May 2015.

[28] R. Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," Oregon Health Sci. Univ., Portland, OR, USA, Tech. Rep., 2004.

[29] Y. Huang, Y. Zhang, N. Li, and L. Zhao, "Improved square-root cubature information filter," *Trans. Inst. Meas. Control*, vol. 39, no. 4, pp. 579–588, Apr. 2017.

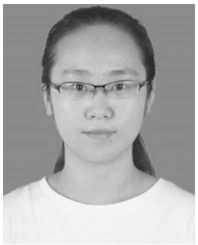
[30] G. Battistelli and L. Chisci, "Kullback–Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707–718, Mar. 2014.

[31] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.

JIAHAO ZHANG is currently pursuing the Ph.D. degree with the Research and Development Institute, Northwestern Polytechnical University in Shenzhen, China. His research interests include control theory and engineering, navigation, guidance and control, target tracking, signal processing, and information fusion.

SHESHENG GAO is currently a Professor with the Research and Development Institute, Northwestern Polytechnical University in Shenzhen, China. His research interests include control theory and engineering, navigation, guidance and control, optimum estimation and control, integrated inertial navigation system, and information fusion.

XIAOMIN QI received the B.Sc. degree in electronic and information engineering from Henan Normal University, in 2010, and the M.Sc. degree in control theory and control engineering from the Zhongyuan University of Technology, in 2014. He is currently pursuing the Ph.D. degree in wireless communication with COMSATS University Islamabad.



JIAHUI YANG is currently pursuing the Ph.D. degree with the Research and Development Institute, Northwestern Polytechnical University in Shenzhen, China. Her research interests include control theory and engineering, navigation, guidance and control, as well as target tracking.

JUAN XIA is currently pursuing the Ph.D. degree with the Research and Development Institute, Northwestern Polytechnical University in Shenzhen, China. Her research interests include control theory and engineering, navigation, guidance and control, optimum estimation and control, information fusion, target tracking, and integrated navigation.

BINGBING GAO received the Ph.D. degree in navigation, guidance, and control from Northwest Polytechnical University, Xi'an, China, in 2018. He is currently an Assistant Professor with the Research and Development Institute, Northwestern Polytechnical University in Shenzhen. His research interests include information fusion, nonlinear filtering, and integrated navigation.