



Brief Paper

Distributed input and state estimation for non-linear discrete-time systems with direct feedthrough

Jinglin Ding¹, Jian Xiao¹, Yong Zhang²

¹School of Electrical Engineering, Southwest Jiaotong University (SWJTU), Chengdu 610031, People's Republic of China

²Electric Power Dispatching Control Center, State Grid Chengdu Electric Power Supply Company, Chengdu 610041, People's Republic of China

E-mail: dingjialing2008@163.com

Abstract: This study investigates the problem of distributed estimation for non-linear system of sensor networks with unknown inputs affecting both the system state and outputs. A novel 'information filtering algorithm' is derived by reconstructing the non-linear version of the extended recursive three-step filter (NERTSF) into the information filter architecture, which simultaneously estimates the state and the unknown input, denoted as non-linear version of the extended recursive three-step information filter (NERTSF). Afterwards the information filter is extended to the 'derivative-free' version with the help of the cubature Kalman filter (CKF) according to the linear error propagation methodology. A distributed filtering algorithm, based on the derivative-free version of the NERTSIF is proposed in which each sensor node only fuses the local observation instead of the global information and updates the local information state and matrix from its neighbours' estimates using the dynamic average-consensus strategy. The efficacy of the proposed distributed algorithm is demonstrated by simulation examples on target tracking problem and is compared with existing algorithms such as centralised fusion filter and distributed CKF, which lack in tracking the true dynamics of the unknown input.

1 Introduction

Large-scale sensor networks are present in many engineering application domains including target tracking, monitoring, robot navigation, road traffic networks, wireless sensor/actuator networks and so on [1–4]. Distributed estimation for sensors requires reduced communication and provides more robustness to sensor failures compared with traditional centralised schemes [5, 6]. Therefore it has attracted considerable attention of researchers in past decades [7–9]. In a distributed framework, filter only need to communicate the information with its neighbouring sensors according to the topology of the given sensor networks and the objective of filtering can be performed in a distributed way. Distributed algorithms of Kalman filtering for discrete-time linear system have been developed by employing dynamic average-consensus strategies [10–12]. Based on the extended Kalman filter (EKF), the method can be directly extended to non-linear Gaussian system [13]. However, EKF based methods may provide poor performance when the non-linearity is severe. Unscented Kalman filter (UKF) provides better filtering accuracy and robustness under the same computational cost, and the distributed version of UKF has been derived by using statistical linear error propagation and introducing the pseudo observation matrix [14, 15]. However, the performance of UKF is highly affected by underlying parameters and CKF [16] which is the direct Gaussian approximation to the Bayesian filter,

and provides more precise state estimation than existing Gaussian filters has been explored for multi-sensor state estimation [17]. A distributed multiple model estimator for simultaneous localisation and tracking (SLAT) with non-line-of-sight (NLOS) mitigation has been proposed in [18], in which the state estimates for each sensor node are derived by the basic interacting multiple model (IMM) approach and the CKF technique. Furthermore, a distributed information filter has been proposed using the consensus filters, in which the information state vectors and information matrices are communicated between sensor nodes by a consensus filter [19, 20]. It is shown that the state estimate is unbiased and the actual covariance matrix is comparable to the centralised fusion filter. The distributed average set-membership filtering of spatially varying processes is investigated for the sensor network in [21], where system under consideration contains sensor saturation in the presence of unknown-but-bounded process and measurement noise rather than stochastic Gaussian white noise in the Kalman filtering.

Nevertheless, the state estimation which is derived from the previously distributed filter is generally based on the assumption that all system parameters and inputs are known. Unfortunately, in many practical applications the systems are often subject to modelling errors and disturbances which are difficult or not possible to measure, and are often treated as stochastic processes with unknown statistics. Thus the performance of this distributed filter may fail in the presence

of model uncertainties and when no prior knowledge about input is available.

On the other hand, the unknown input filtering and state estimation problem for stochastic systems has been intensively investigated ever since the original work of Kitanidis [22]. A common approach to solve this problem is to produce the unknown input decoupled state estimation. By making use of the two-stage Kalman filtering technique and a specific unknown input filtering technique, a robust two-stage Kalman filter which is independent of the underlying input model is derived [23]. However, this is limited to unknown inputs that only affect the systems model. In recent years, many researchers have devoted attention to this more general unknown input filtering problem [23–27]. In [24, 25], the recursive three-step filter (RTSF) was developed on joint input and state estimation to systems with arbitrary unknown inputs in the absence and presence of the direct feedthrough, respectively. However, the RTSF method is applicable only to the case when direct feedthrough matrix (of the unknown input to the output) has full rank. An extension of [26] is addressed for the general case where not only unknown inputs affect both the system state and the output, but also the direct feedthrough matrix has arbitrary rank. However, the global optimality of these filters was not addressed in these results. A globally optimal filtering framework is developed for unbiased minimum-variance state estimation for systems with unknown inputs in [27, 28]. Recently, an unknown input filtering algorithm, named as NERTSF, is proposed to deal with the input and state estimation for non-linear systems [29]. However, this method may not be applicable for the state estimation of some complex non-linear system, since the derivatives of the non-linear system are needed. To compensate these drawbacks, a derivative-free version of the NERTSF (DNERTSF) and its robust two-stage form have been proposed for general non-linear systems with arbitrary unknown inputs in [30, 31]. Most recently, considering the constraints on the states, a non-linear filter based on UKF and the least-squares unbiased estimation algorithm is proposed which simultaneously estimates the constrained state and the unknown input [32]. Besides, a Bayesian framework for simultaneous input and state estimation of non-linear systems with and without input–output direct feedthrough has been developed, and applied to the flow field estimation [33].

It should be noted that the problem filtering with unknown input can also arise in multi-sensor fusion. In a recent work [34], the distributed information consensus filter (ICF) for simultaneous input and state estimation of linear discrete-time system without direct feedthrough in sensor networks is presented. However, to the best of authors' knowledge, the distributed estimator for the simultaneous input and state estimation of non-linear systems with arbitrary unknown inputs has not been investigated yet. Unknown inputs for non-linear systems of sensor networks are inevitable in many practical situations. Moreover, the unknown inputs cannot only influence the system model, it also has impact on outputs.

Owing to the various limitations as discussed above, the focus of this paper is on distributed state and input estimation for discrete-time non-linear systems with unknown inputs affecting both the system and outputs. We extend unknown input filtering problem for non-linear discrete-time systems to the distributed estimation. The main contribution of this paper is the development of a novel distributed filtering for non-linear systems in sensor networks that simultaneously estimates the state and the unknown input, whereas

local information filters are derived for each sensor and the filtering estimates are fused with its neighbours according to the topology of the given sensor networks. Considering that the update stage of information filters is computationally economic and can be easily extended to multi-sensor fusion, the information form of the NERTSF is developed in terms of the inverse of the covariance matrix and the state information estimate, denoted as NERTSIF. Furthermore, the results are extended for derivative-free version of the information filter by combining the CKF algorithm with the statistical linear error propagation methodology. Afterwards, local information filter for each sensor node is derived which only fuses the local observation information. As the average-consensus strategy has been shown to be an effective tool to solve the problem of distributed fusion, this scheme is used to communicate the local information state and matrix with its neighbours, and then the state and the unknown input estimates can be performed in a distributed manner. Simulation results show that the performance of the proposed distributed filter is close to the centralised fusion filter and better than distributed CKF with unknown inputs.

The remainder of this paper is organised as follows. Section 2 presents the problem of distributed estimation for non-linear system with direct feedthrough of the unknown input. Section 3 proposes information filtering algorithm of the NERTSF, labelled as NERTSIF. In Section 4, the derivative-free version of the NERTSIF is developed for the non-linear system with unknown input, and it is further extended to the distributed estimation using an average-consensus filter. Numerical simulations for performance comparison are presented in Section 5, and conclusions are stated in Section 6.

2 Problem formulation

Consider the following discrete-time non-linear system with unknown inputs

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{d}_k, \mathbf{u}_k) + \boldsymbol{\omega}_k \quad (1)$$

$$\mathbf{z}_k^i = \mathbf{g}^i(\mathbf{x}_k, \mathbf{d}_k, \mathbf{u}_k) + \mathbf{v}_k^i, \quad i = 1, 2, \dots, N \quad (2)$$

where $\mathbf{x}_k \in \mathbf{R}^n$ is the state vector, $\mathbf{d}_k \in \mathbf{R}^p$ is an unknown input vector, $\mathbf{u}_k \in \mathbf{R}^q$ is the known input vector and $\mathbf{z}_k^i \in \mathbf{R}^m$ represents the i th sensor measurement vector. N is the number of sensor nodes. \mathbf{f} and \mathbf{g}^i are the system transition function and the measurement function of the i th sensor, respectively. The process noise $\boldsymbol{\omega}_k \in \mathbf{R}^n$ and the measurement noise \mathbf{v}_k^i are assumed to be mutually uncorrelated zero mean white Gaussian noise with known covariance matrices $\mathbf{Q}_k = E[\boldsymbol{\omega}_k(\boldsymbol{\omega}_k)^T] > 0$ and $\mathbf{R}_k^i = E[\mathbf{v}_k^i(\mathbf{v}_k^i)^T] > 0$. We assume that the initial state \mathbf{x}_0 is with unbiased estimate $\hat{\mathbf{x}}_0$ with covariance matrix \mathbf{P}_0 and is independent of $\boldsymbol{\omega}_k$ and \mathbf{v}_k^i . The communication topology between sensor agents is described by the directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, 2, 3, \dots, N\}$ denotes the node set and $\mathcal{E} \subset \{(i, j) | i, j \in \mathcal{V}\}$ denotes the edge set. The set of neighbours of sensor node i is denoted as $\mathcal{N}_i = \{j | (i, j) \in \mathcal{E}\}$. The adjacency matrix $\mathbf{A} = [a_{ij}] \in \mathbf{R}^{N \times N}$ of the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is defined such that $a_{ij} > 0$ if the edge $(i, j) \in \mathcal{E}$, otherwise $a_{ij} = 0$. Moreover, we assume that each node can communicate with itself, so $a_{ii} > 0$ for all $i \in \mathcal{V}$.

Let $\mathbf{z}_k = [\mathbf{z}_k^1; \mathbf{z}_k^2; \dots; \mathbf{z}_k^N] \in \mathbf{R}^{Nm}$ be the central measurement of the entire sensor network at time k . Defining the observation function matrix $\mathbf{g} = [\mathbf{g}^1; \mathbf{g}^2; \dots; \mathbf{g}^N]$, the multi-sensor observation model can be equivalently represented by

the following centralised observation model

$$\mathbf{z}_k = \mathbf{g}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{v}_k \quad (3)$$

Since measurement noise of the sensors \mathbf{v}_k^i is assumed uncorrelated, the global covariance matrix of the measurement noise $\mathbf{v}_k = [\mathbf{v}_k^1; \mathbf{v}_k^2; \dots; \mathbf{v}_k^N]$ can be written as $\mathbf{R}_k = \text{diag}(\mathbf{R}_k^1, \mathbf{R}_k^2, \dots, \mathbf{R}_k^N)$.

The object of this paper is to design a distributed recursive filter which can estimates the system state \mathbf{x}_k and the unknown input \mathbf{d}_k for the discrete-time non-linear system (1) and (2). As prior information about the dynamical evolution of \mathbf{d}_k is assumed unavailable, the distributed filter obtains the simultaneous input and state estimation for each sensor node and the performance of this filter is expected to be comparable to that of the centralised fusion filter which involves the measurements from all sensor nodes.

3 Non-linear version of the extended recursive three-step information filter

This section presents the NERTSIF algorithm, which use NERTSF in an extended information filter framework [13] by restructuring its recursive algorithm in terms of information state vector and the information matrix.

The basic approach of NERTSF has been recently proposed by Hsieh [29]. Using first-order Taylor approximations of (1) and (3), the approximated non-linear system can be obtained as

$$\mathbf{x}_{k+1} \simeq \mathbf{f}(\hat{\mathbf{x}}_{k|k}, \hat{\mathbf{d}}_{k|k}, \mathbf{u}_k) + \mathbf{A}_k \tilde{\mathbf{x}}_{k|k} + \mathbf{G}_k \tilde{\mathbf{d}}_{k|k} + \boldsymbol{\omega}_k \quad (4)$$

$$\begin{aligned} \mathbf{z}_k &\simeq \mathbf{g}(\hat{\mathbf{x}}_{k|k-1}, \hat{\mathbf{d}}_{k-1|k-1}, \mathbf{u}_k) + \mathbf{C}_k \tilde{\mathbf{x}}_{k|k-1} \\ &+ \mathbf{H}_k (\mathbf{d}_k - \hat{\mathbf{d}}_{k-1|k-1}) + \mathbf{v}_k \end{aligned} \quad (5)$$

where $\tilde{\mathbf{x}}_{k|k} \simeq \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}$, $\tilde{\mathbf{d}}_{k|k} \simeq \mathbf{d}_k - \hat{\mathbf{d}}_{k|k}$, $\tilde{\mathbf{x}}_{k|k-1} \simeq \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}$ and

$$\mathbf{A}_k = \left. \frac{\partial \mathbf{f}(\mathbf{x}_k, \mathbf{d}_k, \mathbf{u}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k}, \mathbf{d}_k = \hat{\mathbf{d}}_{k|k}} \quad (6)$$

$$\mathbf{G}_k = \left. \frac{\partial \mathbf{f}(\mathbf{x}_k, \mathbf{d}_k, \mathbf{u}_k)}{\partial \mathbf{d}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k}, \mathbf{d}_k = \hat{\mathbf{d}}_{k|k}} \quad (7)$$

$$\mathbf{C}_k = \left. \frac{\partial \mathbf{g}(\mathbf{x}_k, \mathbf{d}_k, \mathbf{u}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}, \mathbf{d}_k = \hat{\mathbf{d}}_{k-1|k-1}} \quad (8)$$

$$\mathbf{H}_k = \left. \frac{\partial \mathbf{g}(\mathbf{x}_k, \mathbf{d}_k, \mathbf{u}_k)}{\partial \mathbf{d}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}, \mathbf{d}_k = \hat{\mathbf{d}}_{k-1|k-1}} \quad (9)$$

Throughout this paper, we assume that

- (1) The existence condition of the optimal filter $\text{rank } \mathbf{C}_k \mathbf{G}_{k-1} = \text{rank } \mathbf{G}_{k-1} = m$ for all k is satisfied [13].
- (2) $\text{rank } \mathbf{H}_k = m$.

This second assumption is the special case of NERTSF such as the matrix \mathbf{H}_k is of full-column rank, and hence the algorithm can be stated as:

Step 1: Time update

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{f}(\hat{\mathbf{x}}_{k-1|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{u}_{k-1}) \quad (10)$$

(see (11))

Step 2: Estimation of unknown input

$$\tilde{\mathbf{R}}_k = \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T + \mathbf{R}_k \quad (12)$$

$$\mathbf{M}_k = (\mathbf{H}_k^T \tilde{\mathbf{R}}_k^{-1} \mathbf{H}_k)^{-1} \mathbf{H}_k^T \tilde{\mathbf{R}}_k^{-1} \quad (13)$$

$$\hat{\mathbf{d}}_k = \mathbf{M}_k (\mathbf{z}_k - \mathbf{g}^*(\hat{\mathbf{x}}_{k|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{u}_k)) \quad (14)$$

$$\mathbf{P}_{d,k|k} = \mathbf{M}_k \tilde{\mathbf{R}}_k \mathbf{M}_k^T \quad (15)$$

where $\mathbf{g}^*(\hat{\mathbf{x}}_{k|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{u}_k) = \mathbf{g}(\hat{\mathbf{x}}_{k|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{u}_k) - \mathbf{H}_k \hat{\mathbf{d}}_{k-1}$.

Step 3: Measurement update

$$\mathbf{K}_k = \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T \tilde{\mathbf{R}}_k^{-1} \quad (16)$$

$$\mathbf{L}_k = \mathbf{K}_k (\mathbf{I} - \mathbf{H}_k \mathbf{M}_k) \quad (17)$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}) \quad (18)$$

$$\mathbf{P}_{x,k|k} = \mathbf{P}_{x,k|k-1} - \mathbf{L}_k \mathbf{C}_k \mathbf{P}_{x,k|k-1} \quad (19)$$

$$\mathbf{P}_{xd,k|k} = (\mathbf{P}_{dx,k|k})^T = -\mathbf{P}_{x,k|k-1} \mathbf{C}_k^T \mathbf{M}_k^T \quad (20)$$

where $\hat{\mathbf{z}}_{k|k-1} = \mathbf{g}^*(\hat{\mathbf{x}}_{k|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{u}_k)$.

In order to derive the information filtering framework of the NERTSF algorithm, we first rewrite (17) as (see (21))

Defining

$$\bar{\mathbf{R}}_k = (\mathbf{I} - \mathbf{H}_k \mathbf{M}_k)^{-1} \tilde{\mathbf{R}}_k - \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T \quad (22)$$

The gain matrix \mathbf{L}_k can be equivalently expressed as

$$\mathbf{L}_k = \mathbf{P}_{x,k|k}^T \mathbf{C}_k^T \bar{\mathbf{R}}_k^{-1} \quad (23)$$

and the error covariance matrix $\mathbf{P}_{x,k|k}$ can be represented by (see (24))

$$\mathbf{P}_{x,k|k-1} = [\mathbf{A}_{k-1} \quad \mathbf{G}_{k-1}] [\mathbf{P}_{x,k-1|k-1} \quad \mathbf{P}_{xd,k-1|k-1} // \mathbf{P}_{dx,k-1|k-1} \quad \mathbf{P}_{d,k-1|k-1}] [\mathbf{A}_{k-1} \quad \mathbf{G}_{k-1}]^T + \mathbf{Q}_{k-1} \quad (11)$$

$$\begin{aligned} \mathbf{L}_k &= \mathbf{K}_k (\mathbf{I} - \mathbf{H}_k \mathbf{M}_k) \\ &= \mathbf{P}_{x,k|k} (\mathbf{P}_{x,k|k-1} - \mathbf{L}_k \mathbf{C}_k \mathbf{P}_{x,k|k-1})^{-1} \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T \tilde{\mathbf{R}}_k^{-1} (\mathbf{I} - \mathbf{H}_k \mathbf{M}_k) \\ &= \mathbf{P}_{x,k|k} ((\mathbf{I} - \mathbf{H}_k \mathbf{M}_k)^{-1} \tilde{\mathbf{R}}_k (\mathbf{C}_k^T)^{-1} - \mathbf{C}_k \mathbf{P}_{x,k|k-1})^{-1} \\ &= \mathbf{P}_{x,k|k} \mathbf{C}_k^T ((\mathbf{I} - \mathbf{H}_k \mathbf{M}_k)^{-1} \tilde{\mathbf{R}}_k - \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T)^{-1} \end{aligned} \quad (21)$$

$$\begin{aligned} \mathbf{P}_{x,k|k} &= \mathbf{P}_{x,k|k-1} - \mathbf{L}_k \mathbf{C}_k \mathbf{P}_{x,k|k-1} \\ &= \mathbf{P}_{x,k|k-1} - \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T \tilde{\mathbf{R}}_k^{-1} (\mathbf{I} - \mathbf{H}_k \mathbf{M}_k) \mathbf{C}_k \mathbf{P}_{x,k|k-1} \\ &= \mathbf{P}_{x,k|k-1} - \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T ((\mathbf{I} - \mathbf{H}_k \mathbf{M}_k)^{-1} \tilde{\mathbf{R}}_k - \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T + \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T)^{-1} \mathbf{C}_k \mathbf{P}_{x,k|k-1} \\ &= \mathbf{P}_{x,k|k-1} - \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T (\bar{\mathbf{R}}_k + \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T)^{-1} \mathbf{C}_k \mathbf{P}_{x,k|k-1} \end{aligned} \quad (24)$$

Applying the matrix inversion lemma as

$$(A + BD^{-1}C)^{-1} = A^{-1} - A^{-1}B(D + CA^{-1}B)CA^{-1} \quad (25)$$

Then, the error covariance matrix $P_{x,k|k}$ can be rewritten by

$$P_{x,k|k} = [(P_{x,k|k-1})^{-1} + C_k^T \bar{R}_k^{-1} C_k]^{-1} \quad (26)$$

Let us define the Fisher information matrix $Y_{k|k}$ and information state vector $\hat{y}_{k|k}$ as

$$Y_{k|k} = (P_{x,k|k})^{-1} \quad (27)$$

$$\hat{y}_{k|k} = (P_{x,k|k})^{-1} \hat{x}_{k|k} = Y_{k|k} \hat{x}_{k|k} \quad (28)$$

Then, the update equations for the information matrix and the information state vector are obtained as

$$Y_{k|k} = Y_{k|k-1} + I_k \quad (29)$$

$$\begin{aligned} \hat{y}_{k|k} &= \hat{y}_{k|k-1} + C_k^T \bar{R}_k^{-1} (v_k + C_k \hat{x}_{k|k-1}) \\ &= \hat{y}_{k|k-1} + i_k \end{aligned} \quad (30)$$

where i_k is the information state contribution and I_k is its associated information matrix defined by

$$i_k = C_k^T \bar{R}_k^{-1} (v_k + C_k \hat{x}_{k|k-1}) \quad (31)$$

$$I_k = C_k^T \bar{R}_k^{-1} C_k \quad (32)$$

and $v_k = z_k - \hat{z}_{k|k-1}$ is the innovation vector. The predicted information matrix and the predicted information state vector are calculated as

$$Y_{k|k-1} = (P_{x,k|k-1})^{-1} \quad (33)$$

$$\hat{y}_{k|k-1} = (P_{x,k|k-1})^{-1} \hat{x}_{k|k-1} = Y_{k|k-1} \hat{x}_{k|k-1} \quad (34)$$

Hence NERTSIF algorithm can be summarised in Algorithm (Fig. 1).

where N_t denotes the number of total time step.

4 Distributed DNRESTF (DF-DNRESTF)

In this section, the derivative-free version of the NERTSIF is derived from exploiting CKF algorithm and it is extended to develop a distributed filtering for non-linear system (1) and (2) by using the average-consensus theory.

4.1 Derivative-free version of the NERTSIF

As it is stated previously that the NERTSIF is similar to NERTSF that may not be feasible for implementation, the derivatives of the non-linear system model are needed. Then we can extend the NERTSF to a more general derivative-free version utilising the technique of DNERTSF [13]. The main idea of the DNERTSF is to generate a set of sampled points whose mean and variance approximates to the distribution of real state estimation. In addition, considering that CKF is the closest known Gaussian approximation to the Bayesian filter and does not require evaluation of Jacobians during the

estimation process, the focus of this section is to derive the derivative-free version of the NERTSIF with the sampling strategy of the CKF [13].

(1) *Time update stage*: First, we choose the following augmented state vector

$$X_k = [x_k^T \quad d_k^T]^T \quad (35)$$

According to the CKF, the factorisation of the error covariance matrix, evaluation of cubature points and propagated cubature points for the process model is required. Based upon the state estimate $\hat{X}_{k-1|k-1}$ and its covariance $P_{X,k-1|k-1}$

$$\hat{X}_{k-1|k-1} = [\hat{x}_{k-1|k-1}^T \quad \hat{d}_{k-1}^T]^T \quad (36)$$

$$P_{X,k-1|k-1} = \begin{bmatrix} P_{x,k-1|k-1} & P_{xd,k-1|k-1} \\ P_{dx,k-1|k-1} & P_{d,k-1|k-1} \end{bmatrix} \quad (37)$$

The evaluation of an ensemble set of cubature points $\chi_{k-1|k-1}^{(l)}$ ($l = 1, 2, \dots, L$) about the known augmented state estimate can be given as

$$P_{X,k-1|k-1} = S_{k-1|k-1} S_{k-1|k-1}^T \quad (38)$$

$$\chi_{k-1|k-1}^{(l)} = S_{k-1|k-1} \xi_l + \hat{X}_{k-1|k-1} \quad (39)$$

where the notation (l) in $\chi_{k-1|k-1}^{(l)}$ denotes the l th cubature point, $L = 2n_x$, $\xi_l = \sqrt{\frac{L}{2}} [1]_l$ and $[1]_l$ is the l th element of the following set

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -1 \end{pmatrix} \right\}$$

Secondly, evaluating the propagated cubature points

$$x_{k|k-1}^{(l)} = f(x_{k-1|k-1}^{(l)}, u_{k-1}) \quad (40)$$

the predicted state and its associated covariance are given by

$$\hat{x}_{k|k-1} = \frac{1}{L} \sum_{l=1}^L x_{k|k-1}^{(l)} \quad (41)$$

$$P_{x,k|k-1} = \frac{1}{L} \sum_{l=1}^L x_{k|k-1}^{(l)} x_{k|k-1}^{(l)T} - \hat{x}_{k|k-1} \hat{x}_{k|k-1}^T + Q_{k-1} \quad (42)$$

(2) *Measurement update stage*: For measurement update, the propagated cubature points for measurement model can be evaluated as

$$Z_{k|k-1}^{(l)} = g(x_{k-1|k-1}^{(l)}, \hat{d}_{k-1}, u_k) - \hat{H}_k \hat{d}_{k-1} \quad (43)$$

where $\hat{H}_k = [\hat{H}_k^1 \dots \hat{H}_k^p]$ is suitable partial derivative matrix, in which

$$\hat{H}_k^j = \frac{g(\hat{x}_{k|k-1}, \hat{d}_{k-1} + \Delta_j \times e_p^j, u_k) - g(\hat{x}_{k|k-1}, \hat{d}_{k-1}, u_k)}{\Delta_j} \quad (44)$$

Algorithm 1

Input: $\hat{\mathbf{x}}_{0|0}, \mathbf{P}_{x,0|0}, \hat{\mathbf{d}}_0, \mathbf{P}_{d,0|0}, \mathbf{P}_{xd,0|0}$

Output: $\hat{\mathbf{x}}_{k|k}, \hat{\mathbf{d}}_k$

for $k=1,2,\dots,N_t$ **do**

- 1: According to the prior knowledge $(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{x,k|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{P}_{d,k|k-1}, \mathbf{P}_{xd,k|k-1})$, calculate the prediction state $\hat{\mathbf{x}}_{k|k-1}$ and its error covariance matrix $\mathbf{P}_{x,k|k-1}$

$$\hat{\mathbf{x}}_{k|k-1} = f(\hat{\mathbf{x}}_{k-1|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{u}_{k-1})$$

$$\mathbf{P}_{x,k|k-1} = \begin{bmatrix} \mathbf{A}_{k-1} & \mathbf{G}_{k-1} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{x,k-1|k-1} & \mathbf{P}_{xd,k-1|k-1} \\ \mathbf{P}_{dx,k-1|k-1} & \mathbf{P}_{d,k-1|k-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{k-1} & \mathbf{G}_{k-1} \end{bmatrix}^T + \mathbf{Q}_{k-1}$$

- 2: Compute the input estimate $\hat{\mathbf{d}}_k$, its error covariance matrix $\mathbf{P}_{d,k|k-1}$ and the cross-correlation covariance $\mathbf{P}_{xd,k|k-1}$

$$\hat{\mathbf{d}}_k = \mathbf{M}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})$$

$$\mathbf{P}_{d,k|k-1} = \mathbf{M}_k \tilde{\mathbf{R}}_k \mathbf{M}_k^T$$

$$\mathbf{P}_{xd,k|k-1} = (\mathbf{P}_{dx,k|k-1})^T = -\mathbf{P}_{x,k|k-1} \mathbf{C}_k^T \mathbf{M}_k^T$$

where $\hat{\mathbf{z}}_{k|k-1} = \mathbf{g}(\hat{\mathbf{x}}_{k|k-1}, \hat{\mathbf{d}}_{k-1}, \mathbf{u}_k) - \mathbf{H}_k \hat{\mathbf{d}}_{k-1}$

- 3: Evaluate the predicted information matrix and the predicted information state vector:

$$\mathbf{Y}_{k|k-1} = (\mathbf{P}_{x,k|k-1})^{-1}$$

$$\hat{\mathbf{y}}_{k|k-1} = (\mathbf{P}_{x,k|k-1})^{-1} \hat{\mathbf{x}}_{k|k-1} = \mathbf{Y}_{k|k-1} \hat{\mathbf{x}}_{k|k-1}$$

- 4: Compute the information state contribution and its associated information matrix

$$\mathbf{i}_k = \mathbf{C}_k^T \bar{\mathbf{R}}_k^{-1} (\mathbf{v}_k + \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})$$

$$\mathbf{I}_k = \mathbf{C}_k^T \bar{\mathbf{R}}_k^{-1} \mathbf{C}_k$$

where $\bar{\mathbf{R}}_k = (\mathbf{I} - \mathbf{H}_k \mathbf{M}_k)^{-1} \tilde{\mathbf{R}}_k - \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T$

- 5: Update the information matrix and information state vector:

$$\mathbf{Y}_{k|k} = \mathbf{Y}_{k|k-1} + \mathbf{I}_k$$

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{i}_k$$

- 6: Recovery of estimated state and error covariance matrix

$$\mathbf{P}_{k|k} = \mathbf{Y}_{k|k}^{-1}$$

$$\hat{\mathbf{x}}_{k|k} = \mathbf{Y}_{k|k}^{-1} \hat{\mathbf{y}}_{k|k}$$

end for

Fig. 1 Algorithm 1: NERTSIF algorithm

and the predicted measurement can be given as

$$\hat{z}_{k|k-1} = \frac{1}{L} \sum_{l=1}^L \mathbf{Z}_{k|k-1}^{(l)} \quad (45)$$

Thus, $\tilde{\mathbf{R}}_k$ in (12) and the cross-correlation covariance $\mathbf{P}_{xz,k|k-1}$ can be obtained as

$$\tilde{\mathbf{R}}_k = \frac{1}{L} \sum_{l=1}^L (\mathbf{Z}_{k|k-1}^{(l)} - \hat{z}_{k|k-1})(\mathbf{Z}_{k|k-1}^{(l)} - \hat{z}_{k|k-1})^T \quad (46)$$

$$\mathbf{P}_{xz,k|k-1} = \frac{1}{L} \sum_{l=1}^L (\mathbf{x}_{k|k-1}^{(l)} - \hat{\mathbf{x}}_{k|k-1})(\mathbf{Z}_{k|k-1}^{(l)} - \hat{z}_{k|k-1})^T \quad (47)$$

It is noteworthy that the information state contribution and its associated information matrix of the NERTSIF in (31) and (32) are explicit functions of the linearised Jacobian of the measurement model. However, the CKF algorithm does not require the Jacobians for measurement update and hence it cannot be directly used in the NERTSIF framework. To circumvent this problem, we can derive the derivative-free version of the NERTSIF by using the statistical linear error propagation methodology [36]. According to the linear error propagation property, the observation covariance and its cross-correlation covariance can be approximated as

$$\begin{aligned} \tilde{\mathbf{R}}_k &= E\{[\mathbf{z}_k - \hat{z}_{k|k-1}][\mathbf{z}_k - \hat{z}_{k|k-1}]^T\} \\ &\simeq E\{[\mathbf{C}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{v}_k][\mathbf{C}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{v}_k]^T\} \end{aligned} \quad (48)$$

$$\begin{aligned} &= \mathbf{C}_k \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T + \mathbf{R}_k \\ \mathbf{P}_{xz,k|k-1} &= E\{[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}][\mathbf{z}_k - \hat{z}_{k|k-1}]^T\} \\ &\simeq E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})[\mathbf{C}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{v}_k]^T\} \quad (49) \\ &= \mathbf{P}_{x,k|k-1} \mathbf{C}_k^T \end{aligned}$$

and the linearisation measurement matrix \mathbf{C}_k can be obtained as follows

$$\mathbf{C}_k = [(\mathbf{P}_{x,k|k-1})^{-1} \mathbf{P}_{xz,k|k-1}]^T = (\mathbf{P}_{xz,k|k-1})^T (\mathbf{P}_{x,k|k-1})^{-1} \quad (50)$$

Now the information state contribution and its associated information matrix in (31) and (32) can be rewritten as

$$\begin{aligned} \mathbf{i}_k &= (\mathbf{P}_{x,k|k-1})^{-1} \mathbf{P}_{xz,k|k-1} \tilde{\mathbf{R}}_k^{-1} \\ &\times (\mathbf{v}_k + (\mathbf{P}_{xz,k|k-1})^T (\mathbf{P}_{x,k|k-1})^{-1} \hat{\mathbf{x}}_{k|k-1}) \end{aligned} \quad (51)$$

$$\mathbf{I}_k = (\mathbf{P}_{x,k|k-1})^{-1} \mathbf{P}_{xz,k|k-1} \tilde{\mathbf{R}}_k^{-1} (\mathbf{P}_{xz,k|k-1})^T (\mathbf{P}_{x,k|k-1})^{-1} \quad (52)$$

where $\tilde{\mathbf{R}}_k$ is evaluated by

$$\tilde{\mathbf{R}}_k = (\mathbf{I} - \hat{\mathbf{H}}_k \mathbf{M}_k)^{-1} \tilde{\mathbf{R}}_k - (\mathbf{P}_{xz,k|k-1})^T (\mathbf{P}_{x,k|k-1})^{-1} \mathbf{P}_{xz,k|k-1} \quad (53)$$

The updated information state vector and information matrix for the derivative-free version of the NERTSIF can be obtained by using \mathbf{I}_k and \mathbf{i}_k from (51) and (52) in (29) and (30). We can summarise the derivative-free version of the NERTSIF, denoted by DNERTSIF in Algorithm 2 (see Fig. 2).

4.2 Distributed filter of the DNERTSIF

Recalling the update equation of the information matrix and information state vector in (29) and (30), the centralised information filter can be achieved by collecting all the sensor observations at a central location. For the sake of the observation fusion, the global observation information contribution can be written as

$$\mathbf{i} = \sum_{i=1}^N \mathbf{i}_k^i \quad (54)$$

$$\mathbf{I} = \sum_{i=1}^N \mathbf{I}_k^i \quad (55)$$

and each local information state contribution and its associated information matrix at the i sensor site are computed by

$$\begin{aligned} \mathbf{i}_k^i &= (\mathbf{P}_{x,k|k-1}^i)^{-1} \mathbf{P}_{xz,k|k-1}^i (\tilde{\mathbf{R}}_k^i)^{-1} \\ &\times (\mathbf{v}_k^i + (\mathbf{P}_{xz,k|k-1}^i)^T (\mathbf{P}_{x,k|k-1}^i)^{-1} \hat{\mathbf{x}}_{k|k-1}^i) \end{aligned} \quad (56)$$

$$\mathbf{I}_k^i = (\mathbf{P}_{x,k|k-1}^i)^{-1} \mathbf{P}_{xz,k|k-1}^i (\tilde{\mathbf{R}}_k^i)^{-1} (\mathbf{P}_{xz,k|k-1}^i)^T (\mathbf{P}_{x,k|k-1}^i)^{-1} \quad (57)$$

For a fully connected network, the information filter can be implemented in the decentralised version if each node begins a common initial information state, since the observation update in (29) and (30) is additive. However, when the sensor network is not fully connected, issues arise in communicating the observations. Obviously, fully decentralised information filter is no longer equivalent to the optimal centralised version. In distributed information filter framework, the local information filter for each sensor node can be found from a localised form of (29) and (30) as

$$\mathbf{Y}_{k|k}^i = \mathbf{Y}_{k|k-1}^i + \mathbf{i}_k^i \quad (58)$$

$$\hat{\mathbf{y}}_{k|k}^i = \hat{\mathbf{y}}_{k|k-1}^i + \mathbf{i}_k^i \quad (59)$$

In order to estimate the state completely, the local estimation accuracy should be improved by sharing the measurements among the sensor nodes of the network. To address the issue of communicating new observations, the consensus filter can be used. In consensus filter, each sensor node only exchanges the message with its neighbours, and the agents can come into agreement concerning the consensus state asymptotically. In this paper, the information consensus protocol is employed for the distributed estimation in which each node updates its state ξ_i using its neighbours' states according to the following rule [37]

$$\xi_i(\tau + 1) = \beta_{ii}(\tau) \xi_i(\tau) + \sum_{j \in \mathcal{N}_i} \beta_{ij}(\tau) \xi_j(\tau) \quad (60)$$

where $\beta_{ij}(\tau)$ is the linear weight on $\xi_j(\tau)$ at node i and τ is the consensus iteration step. Setting $\sum_{j=1}^N \beta_{ij}(\tau) = 1$, $\beta_{ii}(\tau) > 0$ and $\beta_{ij}(\tau) = 0$ for $j \notin \mathcal{N}_i$. Then the consensus protocol can be further equivalent to

$$\xi_i(\tau + 1) = \xi_i(\tau) + \sum_{j=1}^N \beta_{ij}(\tau) (\xi_j(\tau) - \xi_i(\tau)) \quad (61)$$

In our case, the scheme for choosing $\beta_{ij}(\tau)$ is the Metropolis weights, since it is very simple to compute and well suited

Algorithm 2

Input: $\hat{\mathbf{x}}_{0|0}$, $\mathbf{P}_{x,0|0}$, $\hat{\mathbf{d}}_0$, $\mathbf{P}_{d,0|0}$, $\mathbf{P}_{xd,0|0}$

Output: $\hat{\mathbf{x}}_{k|k}$, $\hat{\mathbf{d}}_k$

for $k=1,2,\dots,N_t$ **do**

Time Update

- 1: Obtain the cubature points $\mathbf{x}_{k-1|k-1}^{(l)}$ ($l=1,2,\dots,L$) according to the posterior probability distribution $N(\hat{\mathbf{X}}_{k-1|k-1}, \mathbf{P}_{x,k-1|k-1})$.

- 2: Evaluate the predicted information matrix and the predicted information state vector:

$$\mathbf{Y}_{k|k-1} = (\mathbf{P}_{x,k|k-1})^{-1}$$

$$\hat{\mathbf{y}}_{k|k-1} = (\mathbf{P}_{x,k|k-1})^{-1} \hat{\mathbf{x}}_{k|k-1} = \mathbf{Y}_{k|k-1} \hat{\mathbf{x}}_{k|k-1}$$

where

$$\hat{\mathbf{x}}_{k|k-1} = \frac{1}{L} \sum_{l=1}^L \mathbf{x}_{k|k-1}^{(l)} = \frac{1}{L} \sum_{l=1}^L f(\mathbf{x}_{k-1|k-1}^{(l)}, \mathbf{u}_{k-1})$$

$$\mathbf{P}_{x,k|k-1} = \frac{1}{L} \sum_{l=1}^L \mathbf{x}_{k|k-1}^{(l)} \mathbf{x}_{k|k-1}^{(l)T} - \hat{\mathbf{x}}_{k|k-1} \hat{\mathbf{x}}_{k|k-1}^T + \mathbf{Q}_{k-1}$$

Measurement Update

- 1: Propagate cubature points $\mathbf{x}_{k|k-1}^{(l)}$ by the nonlinear observation function and get

$$\hat{\mathbf{z}}_{k|k-1} = \frac{1}{L} \sum_{l=1}^L \mathbf{Z}_{k|k-1}^{(l)} = \frac{1}{L} \sum_{l=1}^L \mathbf{g}(\mathbf{x}_{k|k-1}^{(l)}, \mathbf{u}_k)$$

$$\tilde{\mathbf{R}}_k = \frac{1}{L} \sum_{l=1}^L (\mathbf{Z}_{k|k-1}^{(l)} - \hat{\mathbf{z}}_{k|k-1})(\mathbf{Z}_{k|k-1}^{(l)} - \hat{\mathbf{z}}_{k|k-1})^T$$

- 2: Obtain the input estimate $\hat{\mathbf{d}}_k$, its error covariance matrix $\mathbf{P}_{d,k|k}$ and the cross-correlation covariance $\mathbf{P}_{xd,k|k}$

$$\hat{\mathbf{d}}_k = \mathbf{M}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})$$

$$\mathbf{P}_{d,k|k} = \mathbf{M}_k \tilde{\mathbf{R}}_k \mathbf{M}_k^T$$

$$\mathbf{P}_{xd,k|k-1} = (\mathbf{P}_{dx,k|k-1})^T = -\mathbf{P}_{xz,k|k-1} \mathbf{M}_k^T$$

- 3: Compute the information state contribution and its associated information matrix

$$\mathbf{i}_k = (\mathbf{P}_{x,k|k-1})^{-1} \mathbf{P}_{xz,k|k-1} \bar{\mathbf{R}}_k^{-1} (\mathbf{v}_k + (\mathbf{P}_{xz,k|k-1})^T (\mathbf{P}_{x,k|k-1})^{-1} \hat{\mathbf{x}}_{k|k-1})$$

$$\mathbf{I}_k = (\mathbf{P}_{x,k|k-1})^{-1} \mathbf{P}_{xz,k|k-1} \bar{\mathbf{R}}_k^{-1} (\mathbf{P}_{xz,k|k-1})^T (\mathbf{P}_{x,k|k-1})^{-1}$$

where

$$\mathbf{P}_{xz,k|k-1} = \frac{1}{L} \sum_{l=1}^L (\mathbf{x}_{k|k-1}^{(l)} - \hat{\mathbf{x}}_{k|k-1})(\mathbf{Z}_{k|k-1}^{(l)} - \hat{\mathbf{z}}_{k|k-1})^T$$

- 4: Update the information matrix and information state vector:

$$\mathbf{Y}_{k|k} = \mathbf{Y}_{k|k-1} + \mathbf{I}_k$$

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{i}_k$$

Recovery of estimated state and error covariance matrix

$$\mathbf{P}_{k|k} = \mathbf{Y}_{k|k}^{-1}$$

$$\hat{\mathbf{x}}_{k|k} = \mathbf{Y}_{k|k}^{-1} \hat{\mathbf{y}}_{k|k}$$

end for

Fig. 2 Algorithm 2: Derivative-free version of the NERTSIF algorithm

for distributed implementation [38]. In particular, each node only needs to know their neighbours' degree to determine the weights.

For the ICF [19], we assume that communication and prediction update are synchronised in the sensor network, each sensor node i updates its local information state vector and matrix from its neighbours' estimates according to the consensus filter as (61) rather than communicating the measurement directly. In the measurement update, each agent i only uses the local observation rather than the global information. For clarity, Algorithm 3 (see Fig. 3) describes the distributed filter of the DNERTSIF for the non-linear system (1) and (2) with the unknown input based on the ICF where T_p is the number of consensus iteration steps to achieve average-consensus.

Remark 1: In Algorithm 3 (Fig. 3), τ is the time index for the consensus protocol and one time step $k-1 \rightarrow k$ is equivalent to T_p iteration steps for the consensus. It means that in the iterative process of each loop for consensus filter, there is only one update of the local information filter which runs at every first step of the consensus iteration.

Remark 2: As shown in [19], the ICF estimate is unbiased and the true variance of the estimate is comparable to the centralised estimate. Then these properties are maintained in the distributed filter of the DNERTSIF for non-linear system with unknown input.

5 Simulation results

5.1 Simulation scenario

In order to verify the performance of the proposed distributed filter, we will consider a two-dimensional radar tracking scenario to compare the performance of the proposed distributed information filter with that of the centralised confusion filter. The target executes turn in an $x-y$ plan with an unknown turn rate ω , and the target dynamics can be described by the following non-linear motion model [15]

$$\mathbf{x}_k = \begin{pmatrix} 1 & \frac{\sin \omega_{k-1} T}{\omega_{k-1}} & 0 & -\left(\frac{1 - \cos \omega_{k-1} T}{\omega_{k-1}} \right) \\ 0 & \cos \omega_{k-1} T & 0 & -\sin \omega_{k-1} T \\ 0 & \frac{1 - \cos \omega_{k-1} T}{\omega_{k-1}} & 1 & \frac{\sin \omega_{k-1} T}{\omega_{k-1}} \\ 0 & \sin \omega_{k-1} T & 0 & \cos \omega_{k-1} T \end{pmatrix} \mathbf{x}_{k-1} + \begin{pmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{pmatrix} \mathbf{w}_{k-1} \quad (62)$$

where the state of the target $\mathbf{x} = [\xi \ \dot{\xi} \ \eta \ \dot{\eta}]^T$; ξ and η represent the target position, $\dot{\xi}$ and $\dot{\eta}$ represent the target velocity in the x and y directions, respectively; ω represents an unknown time-varying input and we assume that there is no prior knowledge about ω is available in the filtering process. T is the sampling time period, the process noise $\mathbf{w}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k-1})$.

Using 12 radars which are fixed at the origin of the plan to measure the range and bearing of target, then the measurement equation can be given by

$$\mathbf{z}_{i,k} = \begin{pmatrix} r_{i,k} \\ \theta_{i,k} \end{pmatrix} = \begin{pmatrix} \sqrt{(\xi_k - x_{i,0})^2 + (\eta_k - y_{i,0})^2} \\ \tan^{-1}((\eta_k - y_{i,0})/(\xi_k - x_{i,0})) \end{pmatrix} + \mathbf{v}_{i,k}, \quad (i = 1, 2, \dots, 12) \quad (63)$$

where $(x_{i,0}, y_{i,0})$ denotes the position of the i th radar sensor, and the measurement noise $\mathbf{v}_{i,k} \sim (\mathbf{0}, \mathbf{R}_{i,k})$ with $\mathbf{R}_{i,k} = \text{diag}[\sigma_r^2 \ \sigma_\theta^2]$ for all sensors in the following simulations. To show the effectiveness and feasibility of the algorithm that simultaneously estimates the state and input for non-linear system of sensor networks with unknown inputs affecting both the system state and outputs, we slightly modify the above measurement model by adding a new measurement which is associated with the unknown input ω_k as follows

$$\mathbf{z}_{i,k} = \begin{pmatrix} r_{i,k} \\ \theta_{i,k} \\ \phi_k \end{pmatrix} = \begin{pmatrix} \sqrt{(\xi_k - x_{i,0})^2 + (\eta_k - y_{i,0})^2} \\ \tan^{-1}((\eta_k - y_{i,0})/(\xi_k - x_{i,0})) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 10 \end{pmatrix} \omega_k + \mathbf{v}_{i,k}, \quad (i = 1, 2, \dots, 12) \quad (64)$$

Comparing (64) with (63), it is clearly that just the measurement $\mathbf{z}_{i,k}$ has an additional component after adding a new variable and the additional variable essentially does not affect the original target dynamics. The corresponding measurement noise with covariance $\mathbf{R}_{i,k} = \text{diag}[\sigma_r^2 \ \sigma_\theta^2 \ \sigma_\phi^2]$ for all sensors and the parameters are given as

$$T = 1 \text{ s}, \quad \mathbf{Q}_{k-1} = 0.04 \mathbf{I}_2, \quad \sigma_r = 0.2 \text{ m}, \\ \sigma_\theta = 0.015 \text{ rad}, \quad \sigma_\phi = 0.017^\circ \text{ s}^{-1}$$

In this tracking scenario, the target executes with different turn rates. First, it makes a coordinate turn motion with turn rate $\omega = -3^\circ \text{ s}^{-1}$ between 1 and 30 s, then makes a coordinate turn motion with turn rate $\omega = 6^\circ \text{ s}^{-1}$ between 31 and 60 s; and at last it makes a coordinate turn motion with turn rate $\omega = 10^\circ \text{ s}^{-1}$ for another 40 s. The true initial state for the target trajectory and its associated covariance are given by

$$\mathbf{x}_0 = [-40 \text{ m} \quad 3 \text{ ms}^{-1} \quad 10 \text{ m} \quad 1 \text{ ms}^{-1}]^T \\ \mathbf{P}_0 = \text{diag}[2^2 \text{ m}^2 \quad 0.1^2 \text{ m}^2 \text{ s}^{-2} \quad 2^2 \text{ m}^2 \quad 0.1^2 \text{ m}^2 \text{ s}^{-2}]$$

The initial estimate state for the filter $\hat{\mathbf{x}}_{0|0}$ is chosen randomly from $\mathcal{N}(\mathbf{x}_0, \mathbf{P}_0)$ and the initial estimate state of the unknown input $\hat{\omega}_0$ is taken as -3° s^{-1} . The distribution of the 12 sensors are shown in Fig. 4, the coordinates are given as $(-20, -40)$, $(20, -40)$, $(60, -40)$, $(100, -40)$, $(-20, -10)$, $(20, -10)$, $(60, -10)$, $(100, -10)$, $(-20, 20)$, $(20, 20)$, $(60, 20)$ and $(100, 20)$, respectively. The communication topology diagram between the sensors is shown in Fig. 5.

To evaluate the performance of the proposed filter, the root-mean-square error (RMSE) in position is used,

Algorithm 3

Input: $\hat{\mathbf{x}}_{0|0}^i$, $\mathbf{P}_{x,0|0}^i$, $\hat{\mathbf{d}}_0^i$, $\mathbf{P}_{d,0|0}^i$, $\mathbf{P}_{xd,0|0}^i$

Output: $\hat{\mathbf{x}}_{k|k}^i, \hat{\mathbf{d}}_k^i$

for $k=1,2,\dots,N_t$ **do**

Time Update

1: Obtain the cubature points $\mathbf{x}_{k-1|k-1}^{i(l)}$ ($l=1,2,\dots,L$) according to the posterior probability

distribution $N(\hat{\mathbf{x}}_{k-1|k-1}^i, \mathbf{P}_{x,k-1|k-1}^i)$.

2: Evaluate the predicted information matrix and the predicted information state vector:

$$\mathbf{Y}_{k|k-1}^i = (\mathbf{P}_{x,k|k-1}^i)^{-1}$$

$$\hat{\mathbf{y}}_{k|k-1}^i = (\mathbf{P}_{x,k|k-1}^i)^{-1} \hat{\mathbf{x}}_{k|k-1}^i = \mathbf{Y}_{k|k-1}^i \hat{\mathbf{x}}_{k|k-1}^i$$

where

$$\hat{\mathbf{x}}_{k|k-1}^i = \frac{1}{L} \sum_{l=1}^L \mathbf{x}_{k|k-1}^{i(l)} = \frac{1}{L} \sum_{l=1}^L f(\mathbf{x}_{k-1|k-1}^{i(l)}, \mathbf{u}_{k-1})$$

$$\mathbf{P}_{x,k|k-1}^i = \frac{1}{L} \sum_{l=1}^L \mathbf{x}_{k|k-1}^{i(l)} \mathbf{x}_{k|k-1}^{i(l)T} - \hat{\mathbf{x}}_{k|k-1}^i (\hat{\mathbf{x}}_{k|k-1}^i)^T + \mathbf{Q}_{k-1}$$

Measurement Update

1: Propagate cubature points $\mathbf{x}_{k|k-1}^{i(l)}$ by the nonlinear observation function and get

$$\hat{\mathbf{z}}_{k|k-1}^i = \frac{1}{L} \sum_{l=1}^L \mathbf{z}_{k|k-1}^{i(l)} = \frac{1}{L} \sum_{l=1}^L \mathbf{g}(\mathbf{x}_{k|k-1}^{i(l)}, \mathbf{u}_k)$$

$$\tilde{\mathbf{R}}_k^i = \frac{1}{L} \sum_{l=1}^L (\mathbf{z}_{k|k-1}^{i(l)} - \hat{\mathbf{z}}_{k|k-1}^i)(\mathbf{z}_{k|k-1}^{i(l)} - \hat{\mathbf{z}}_{k|k-1}^i)^T$$

2: Obtain the input estimate $\hat{\mathbf{d}}_k^i$, its error covariance matrix $\mathbf{P}_{d,k|k}^i$ and the

cross-correlation covariance $\mathbf{P}_{xd,k|k}^i$

$$\hat{\mathbf{d}}_k^i = \mathbf{M}_k^i (\mathbf{z}_k^i - \hat{\mathbf{z}}_{k|k-1}^i)$$

$$\mathbf{P}_{d,k|k}^i = \mathbf{M}_k^i \tilde{\mathbf{R}}_k^i (\mathbf{M}_k^i)^T$$

$$\mathbf{P}_{xd,k|k-1}^i = (\mathbf{P}_{dx,k|k-1}^i)^T = -\mathbf{P}_{xz,k|k-1}^i (\mathbf{M}_k^i)^T$$

3: Compute the local information state contribution and its associated information matrix

$$\mathbf{i}_k^i = (\mathbf{P}_{x,k|k-1}^i)^{-1} \mathbf{P}_{xz,k|k-1}^i (\tilde{\mathbf{R}}_k^i)^{-1} (\mathbf{u}_k^i + (\mathbf{P}_{xz,k|k-1}^i)^T (\mathbf{P}_{x,k|k-1}^i)^{-1} \hat{\mathbf{x}}_{k|k-1}^i)$$

$$\mathbf{I}_k^i = (\mathbf{P}_{x,k|k-1}^i)^{-1} \mathbf{P}_{xz,k|k-1}^i (\tilde{\mathbf{R}}_k^i)^{-1} (\mathbf{P}_{xz,k|k-1}^i)^T (\mathbf{P}_{x,k|k-1}^i)^{-1}$$

Consensus update:

1: Initialisation:

$$\hat{\mathbf{y}}_{k|k-1}^i = \hat{\mathbf{y}}_i[0] \quad \mathbf{Y}_{k|k-1}^i = \mathbf{Y}_i[0]$$

for $\tau=1,2,\dots,T_p$ **do**

$$\hat{\mathbf{y}}_i(\tau) = \hat{\mathbf{y}}_i(\tau-1) + \sum_{j=1}^N \beta_{ij}(\tau) (\hat{\mathbf{y}}_j(\tau-1) - \hat{\mathbf{y}}_i(\tau-1))$$

$$\mathbf{Y}_i(\tau) = \mathbf{Y}_i(\tau-1) + \sum_{j=1}^N \beta_{ij}(\tau) (\mathbf{Y}_j(\tau-1) - \mathbf{Y}_i(\tau-1))$$

2: **if** new observations are taken

then update the local information filter

$$\hat{\mathbf{y}}_i(\tau) \leftarrow \hat{\mathbf{y}}_i(\tau) + \mathbf{i}_{i,k}$$

$$\mathbf{Y}_i(\tau) \leftarrow \mathbf{Y}_i(\tau) + \mathbf{I}_{i,k}$$

end for

3: Obtain the consensus state and its information matrix

$$\hat{\mathbf{y}}_{k|k}^i = \hat{\mathbf{y}}_i[T_p] \quad \mathbf{Y}_{k|k}^i = \mathbf{Y}_i[T_p]$$

Recovery of estimated state and error covariance matrix

$$\mathbf{P}_{k|k}^i = (\mathbf{Y}_{k|k}^i)^{-1}$$

$$\hat{\mathbf{x}}_{k|k}^i = (\mathbf{Y}_{k|k}^i)^{-1} \hat{\mathbf{y}}_{k|k}^i$$

end for

Fig. 3 Algorithm 3: Distributed filter of the DNERTSIF

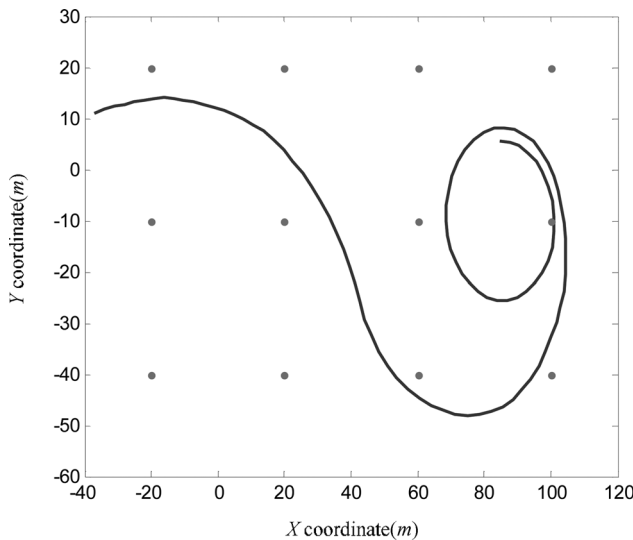


Fig. 4 Target tracking scenario with 12 sensors

defined by

$$\text{RMSE}_k^{\text{pos}} = \left[\frac{1}{M} \sum_{i=1}^M ((\xi_k^i - \hat{\xi}_k^i)^2 + (\eta_k^i - \hat{\eta}_k^i)^2) \right]^{1/2}, \quad (k = 1, 2, \dots, N_t) \quad (65)$$

where (ξ_k^i, η_k^i) and $(\hat{\xi}_k^i, \hat{\eta}_k^i)$ are the true and the estimated position at time k . M is the Monte Carlo runs. Similarly, the formula of the RMSE in velocity may also be written. In the following experiments, we analyse quantitatively and compare the performance of the proposed algorithm according to the RMSE in the position and velocity. Making 50 independent Monte Carlo runs with the same condition and the simulation step of per run N_t take 100. To implement the distributed filter, $T_p = 45$ iteration steps have been used to achieve the average-consensus.

5.2 Simulation results

In the following simulation experiment, the centralised fusion filter which involves the measurement from all the sensor nodes can be used as the baseline algorithm and the proposed distributed filter is compared with it. Furthermore, for emphasising the performance of the proposed distributed filter in simultaneous input and state estimation, the tracking performance using the distributed CKF which adopts the multi-sensor CIF form [17] and the consensus protocol in this paper is also presented. For the sake of simplicity, the centralised fusion filter is denoted by CF-DNERTSF. The distributed CKF and the distributed filter of the DNERTSIF using the i th sensor measurement are shortly denoted by CF-CKFi and DF-DNERTSIFi, respectively.

Fig. 6 shows that the time-varying turn rate ω can be tracked accurately by the estimated unknown input in DF-DNERTSIFi algorithm. Furthermore, the filtering effect of the DF-DNERTSIFi is almost identical and very close to the CF-DNERTSF. It means that the consensus has been reached for each local filter, and also suggests that the performance of the distributed filter of the DNERTSIF is in close agreement with that of the centralised fusion filter.

In order to illustrate the effectiveness of the proposed distributed filter more completely, the performance comparison with respect to RMSE in position and velocity for the above

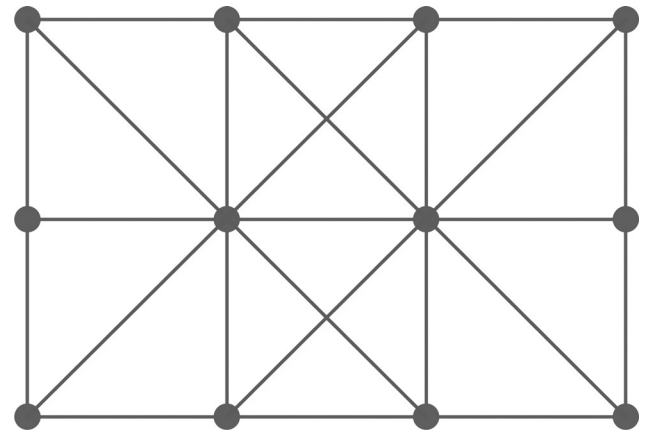


Fig. 5 Communication topology between sensors

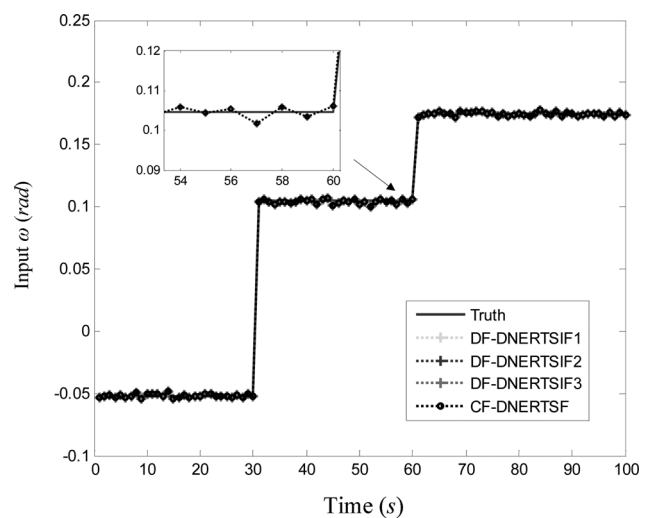


Fig. 6 Actual and estimated input

One sample run

three algorithms is shown in Figs. 7 and 8. It can be observed that the RMSE of the DF-DNERTSIFi is close to that of the centralised fusion filter and much smaller than that of the distributed CKF both in position and in velocity. Making a detailed analysis of the RMSE for the distributed CKF, we observe that it can accurately estimate in initial 10s because of the exact value of the turn rate ω . However, the RMSE start suddenly increased from 30s since the distributed CKF cannot accurately estimate the time-varying turn rate ω . Conversely, the distributed filter of the DNERTSIF can always keep a good filtering performance. This is expected since that the distributed filter of the DNERTSIF provides accurate and robust estimation not only for the system states but also the unknown inputs. Table 1 describes the average RMSE (ARMSE) and its variance (DRMSE) for the three algorithms, which are defined as the average and the variance of RMSE over all time instants in position and velocity, respectively. From Table 1, it can be seen that the performance of the proposed distributed filter is comparable to the centralised fusion filter and superior to the distributed CKF in terms of the filtering precision and stability.

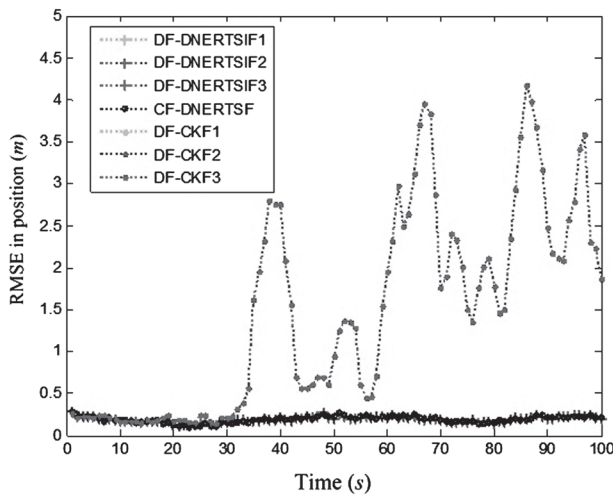


Fig. 7 Performance comparison with respect to RMSE in position

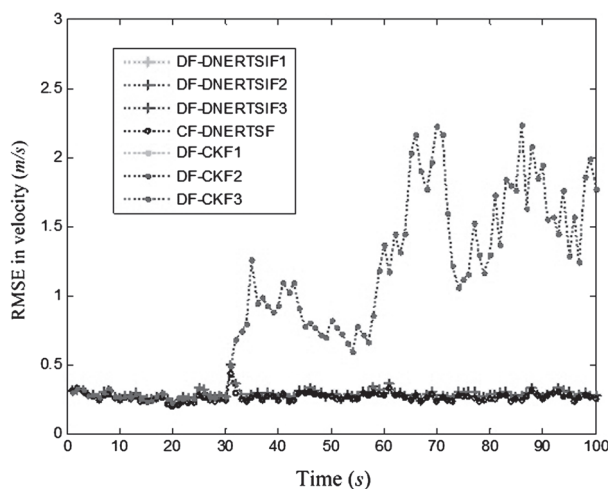


Fig. 8 Performance comparison with respect to RMSE in velocity

Table 1 Performance comparison with respect to AMRSE and DMRSE

Algorithm	AMRSE (position)	AMRSE (velocity)	DMRSE (position)	DMRSE (velocity)
DF-CKF	1.4471	0.9902	1.4237	0.3802
DF-DNERTSIF	0.1905	0.2866	0.0012	0.0011
CF-DNERTSF	0.1900	0.2631	0.0015	0.0009

6 Conclusions

In this paper, a new distributed filter has been developed for the problem of the simultaneous input and state estimation of non-linear system with direct feedthrough of the unknown input in sensor networks. The proposed filter is derived from the derivative-free version of the NERTSIF, in which the local information filtering is derived for each sensor and utilising the average-consensus strategy to fuse the local information state vectors and information matrices between neighbouring sensor nodes. Specifically, the state estimate is unbiased and the actual covariance matrix is close to that of the centralised fusion filter. On the basis of the information filter framework of the NERTSF, the

derivative-free implementation of the NERTSIF is derived for general non-linear system with unknown inputs by applying the CKF algorithm. Simulation results show that the performance of the proposed distributed filter is comparable to the centralised fusion filter which involves all the sensor measurements. Furthermore, this distributed filter provides more accurate and robust estimation than that of the distributed CKF since the latter lacks true dynamic information of the unknown inputs.

In future investigations, the proposed distributed filter could be considered in some practical applications, such as fault detection and isolation for non-linear system with unknown input.

7 Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 51177137, 61134001).

8 References

- Ren, W., Beard, R.W., Atkins, E.M.: 'Information consensus in multi-vehicle cooperative control', *IEEE Control Syst. Mag.*, 2007, **27**, (2), pp. 71–82
- Khan, U.A., Moura, J.M.F.: 'Distributing the Kalman filter for large scale systems', *IEEE Trans. Signal Process.*, 2008, **56**, (10), pp. 4919–4935
- Negenborn, R.R., De Schutter, B., Hellendoorn, J.: 'Multi-agent model predictive control for transportation networks: serial versus parallel schemes', *Eng. Appl. Artif. Intell.*, 2008, **21**, (3), pp. 353–366
- Yu, W., Chen, G., Wang, Z., Yang, W.: 'Distributed consensus filtering in sensor networks', *IEEE Trans. Syst. Man Cybern. B, Cybern.*, 2009, **39**, (6), pp. 1568–1577
- Carli, R., Chiuso, A., Schenato L. *et al.*: 'Distributed Kalman filtering based on consensus strategies', *IEEE J. Sel. Areas Commun.*, 2008, **26**, (4), pp. 622–633
- Olfati-Saber, R., Fax, J.A., Murray, R.M.: 'Consensus and cooperation in networked multi-agents systems', *Proc. IEEE*, 2007, **95**, (1), pp. 215–233
- Sims, C., Park, Y.: 'Ordered sequential filtering for finite dimensional systems with distributed observations', *IEEE Trans. Autom. Control*, 1975, **20**, (3), pp. 351–35
- Alouani, A.T., Gray, J.E., McCabe, D.H.: 'Theory of distributed estimation using multiple asynchronous sensors', *IEEE Trans. Aerosp. Electron. Syst.*, 2005, **41**, (2), pp. 717–722
- Shen, B., Wang, Z.D., Hung, Y.S.: 'Distributed H_∞ -consensus filtering in sensor networks with multiple missing measurements: the finite-horizon case', *Automatica*, 2010, **46**, (10), pp. 1682–1688
- Olfati-Saber, R.: 'Distributed Kalman filter with embedded consensus filters'. Proc. 44th IEEE Conf. Decision and Control and the European Control Conf., Seville, Spain, December 2005, pp. 8179–8184
- Olfati-Saber, R., Shamma, J.: 'Consensus filters for sensor networks and distributed sensor fusion'. Proc. 44th IEEE Conf. Decision and Control and the European Control Conf., Seville, Spain, December 2005, pp. 6698–6703
- Olfati-Saber, R.: 'Distributed Kalman filtering for sensor networks'. Proc. 46th IEEE Conf. Decision and Control, New Orleans, LA, USA, December 2007, pp. 5492–5498
- Mutambara, A.G.O.: 'Decentralized estimation and control for multi-sensor systems' (CRC press, 1998)
- Li, W.L., Jia, Y.M.: 'Distributed consensus filtering for discrete-time nonlinear systems with non-Gaussian noise', *Signal Process.*, 2012, **92**, (10), pp. 2464–2470
- Li, W.L., Jia, Y.M.: 'Consensus-based distributed multiple model UKF for jump Markov nonlinear systems', *IEEE Trans. Autom. Control*, 2012, **57**, (1), pp. 230–236
- Arasaratnam, I., Haykin, S.: 'Cubature Kalman filters', *IEEE Trans. Autom. Control*, 2009, **54**, (6), pp. 1254–1269
- Pakki, K., Chandra, B., Gu, D.W. *et al.*: 'A square root Cubature information filter', *IEEE Sensors J.*, 2013, **13**, (2), pp. 750–758

- 18 Li, W.L., Jia, Y.M., Du J.P. *et al.*: 'Distributed multiple model estimation for simultaneous localization and tracking with NLOS mitigation', *IEEE Trans. Veh. Technol.*, 2013, **62**, (6), pp. 2824–2830
- 19 Casbeer, D.W., Beard, R.: 'Distributed information filtering using consensus filters'. Proc. American Control Conf., Hyatt Regency Riverfront, St. Louis, MO, USA, June 2009, pp. 1882–1887
- 20 Li, W.L., Jia, Y.M.: 'Consensus-based distributed information filter for a class of jump Markov systems', *IET Control Theory Appl.*, 2011, **5**, (10), pp. 1214–1222
- 21 Zhang, H., Yan, H., Yang, F. *et al.*: 'Distributed average filtering for sensor networks with sensor saturation', *IEE Proc., Control Theory Appl.*, 2013, **7**, (6), pp. 1751–8644
- 22 Kitanidis, P.K.: 'Unbiased minimum-variance linear state estimation', *Automatica*, 1987, **23**, (6), pp. 775–778
- 23 Hsieh, C.S.: 'Robust two-stage Kalman filters for systems with unknown inputs', *IEEE Trans. Autom. Control*, 2000, **45**, (12), pp. 2374–2378
- 24 Gillijns, S., De Moor, B.: 'Unbiased minimum-variance input and state estimation for linear discrete-time systems', *Automatica*, 2007, **43**, (1), pp. 111–116
- 25 Gillijns, S., De Moor, B.: 'Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough', *Automatica*, 2007, **43**, (5), pp. 934–937
- 26 Hsieh, C.S.: 'Extension of unbiased minimum-variance input and state estimation for systems with unknown inputs', *Automatica*, 2009, **45**, (9), pp. 2149–2153
- 27 Hsieh, C.S.: 'On the global optimality of unbiased minimum-variance state estimation for systems with unknown inputs', *Automatica*, 2010, **46**, (4), pp. 708–715
- 28 Hsieh, C.S.: 'On the optimality of two-stage Kalman filtering for systems with unknown inputs', *Asian Journal of Control*, 2010, **12**, (4), pp. 510–523
- 29 Hsieh, C.S., Liaw, D.C.: 'Traffic state estimation within unknown input filtering framework'. Proc. 2011 Int. Conf. Service and Interactive Robots, 2011, pp. 25–30
- 30 Hsieh, C.S.: 'State estimation for nonlinear systems with unknown inputs'. Proc. Seventh IEEE Conf. Industrial Electronics and Applications, Singapore, July 2012, pp. 1533–1538
- 31 Hsieh, C.S., Liaw, D.C.: 'Unknown input filtering for nonlinear systems and its application to traffic state estimation'. Proc. Seventh IEEE Conf. Industrial Electronics and Applications, Singapore, July 2012, pp. 1847–1852
- 32 Luo, Z., Fang, H., Luo, Y.: 'Constrained state estimation for nonlinear systems with unknown input', *Circuits Syst. Signal Process.*, 2013, **32**, (5), pp. 2199–2211
- 33 Fang, H., de Callafon, R.A., Cortés, J.: 'Simultaneous input and state estimation for nonlinear systems with applications to flow field estimation', *Automatica*, 2013, **49**, (9), pp. 2805–2812
- 34 Lu, Y., Zhang, L., Mao, X.: 'Distributed information consensus filters for simultaneous input and state estimation', *Circuits Syst. Signal Process.*, 2013, **32**, (2), pp. 877–888
- 35 Darouach, M., Zasadzinski, M.: 'Unbiased minimum variance estimation for systems with unknown exogenous inputs', *Automatica*, 1997, **33**, (4), pp. 717–719
- 36 Lee, D.J.: 'Nonlinear estimation and multiple sensor fusion using unscented information filtering', *IEEE Signal Process. Lett.*, 2008, **15**, pp. 861–864
- 37 Kingston, D., Beard, R.: 'Discrete-time average-consensus under switching network topologies'. Proc. American Control Conf., Minneapolis, MN, USA, June 2006, pp. 3551–3556
- 38 Xiao, L., Boyd, S., Lall, S.: 'A scheme for robust distributed sensor fusion based on average consensus'. Fourth Int. Symp. Information Processing in Sensor Networks, April 2005, pp. 63–70